Universidade de Aveiro Departamento de Economia, Gestão e Engenharia Industrial
2009

Joana Isabel Louro Miguéns

Redes em Turismo: de Destinos Internacionais a Inter-Organizações

Networked Tourism: From World Destinations to Inter-Organizations

Universidade de Aveiro Departamento de Economia, Gestão e Engenharia Industrial
2009

Joana Isabel Louro Miguéns

Redes em Turismo: de Destinos Internacionais a Inter-Organizações

Networked Tourism: From World Destinations to Inter-Organizations

dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Turismo, realizada sob a orientação científica do Professor Doutor Carlos Manuel Martins da Costa, Professor Associado com Agregação do Departamento de Economia, Gestão e Engenharia Industrial da Universidade de Aveiro e da co-orientação do Professor Doutor José Fernando Mendes, Professor Catedrático do Departamento de Física.

Apoio financeiro da FCT
(SFRH / BD / 19258 / 2004) no âmbito do POCI 2010 - Formação Avançada para a Ciência - Medida IV.3.

to Tom Süßenbach

o júri

Presidente

Professora Doutora Nilza Maria Vilhena Nunes da Costa
professor catedrático da Universidade de Aveiro

Professor Doutor Henrique Manuel Morais Diz
professor catedrático da Universidade de Aveiro

Professor Doutor José Fernando Ferreira Mendes
professor catedrático da Universidade de Aveiro

Professor Doutor Carlos Manuel Martins da Costa
professor associado com agregação da Universidade de Aveiro

Professor Doutor Romualdo Pastor Satorras
professor titular do departamento de Física e Enginyeria Nuclear da Universitat Politécnica de Catalunya – Espanha

Professor Doutor Francisco José Lopes de Sousa Diniz
professor associado da Universidade de Trás-os-Montes e Alto Douro

acknowledgements

The thesis is a result of a PhD that started in the University of Aveiro on 2005. My interest and dedication to tourism research lead me to my advisor Professor Carlos Manuel Martins da Costa, and to his researchers' team on tourism in Portugal, to whom I thank their affectionate reception, Carlos Costa, José Manuel Brás, Isabel Martins, Adriana Corfu, Zélia Breda, Emese Panyik, Carla Melo and Rui Costa.

My thesis started co-advised by Professor Alexander Plakhov, to whom I show gratitude for his help and understanding the need to change topic and consequently co-advisor. The work developed to the topic of complex networks which led me to my co-advisor Professor José Fernando F. Mendes, which I thank for stimulating a scientific keenness on myself.

I would like to thank the group of complex networks from the Physics department with whom I shared fruitful seminars, conferences and ideas, Sergey Dorogovtsev, Sasha Goltsev, Sasha Samukhin, João Gama Oliveira, Massimo Ostilli and Zhang Peng.

During my PhD I did a research visit to the group of Luís Amaral in Northwestern University, USA. I thank Luís and all his team Roger Guimerà, Marta Sales-Pardo, Michael Stringer, Dean Malmgren, Daniel Stouffer and Sam Stouffer. Also the support of Andreia Castro and Susana Garcia, and to my friend Jinny Suh.

To my friends with whom I shared my time in Aveiro, Ana Cristina, Maria de Lurdes, Sofia, Ana Filipa, and to Patrícia Engrácia.

During my first year of doctoral PhD research I visited Professor João Abreu de Faria Bilhim whom I also thank for enthusiastically introducing organizational theory to myself.

The final and most important special acknowledgement to my family Maria Manuela Louro, José Miguéns, Maria de Almeida Miguéns, Pedro Miguéns, also to Susana Janeiro and to Tom Süßenbach, for their treasured support.

palavras-chave

Redes em turismo, teoria de redes sociais, crescimento de fluxos em turismo, teoria de grafos, redes complexas.

resumo

Na última década, a referência ao conceito de redes cresceu rapidamente entre a literatura sobre turismo, geralmente aplicado a tópicos como as inter-organizações, estrutura de multi-destinos, espaços de Turismo online, entre outros.

O conceito de rede difundiu-se na natureza e na sociedade, em áreas que vão desde a Biologia à Medicina, ou da Economia à Gestão, e o conhecimento sobre redes tem vindo a impulsionar uma teoria comum para facilitar a compreensão de diferentes sistemas complexos e a representação das ligações entre organizações, acções, bens, proteínas ou pessoas.

A tese teve como propósito o encontro de um eixo comum entre dois campos férteis de investigação através de uma revisão teórica sistemática. A investigação sobre redes complexas é um campo recente na Física que tem vindo a desenvolver-se bastante na última década com fortes aplicações interdisciplinares. Por outro lado, a análise de redes sociais é uma área de investigação activa em Sociologia e Economia há bastante tempo. O estudo das implicações das redes complexas para a ciência das redes de turismo é uma área promissora já com resultados fascinantes.

A tese tem três resultados principais. Primeiro, traz conhecimento das ricas áreas de conhecimento sobre redes complexas e redes sociais. Em segundo lugar, apresenta modelos evolutivos que melhor se adaptam às chegadas turísticas internacionais. Como se organizam as redes sociais? Como é que os indivíduos escolhem os seus destinos de viagem? Estes são exemplos de questões que serão abordadas na tese.

Em terceiro lugar, discute resultados que fazem notar comportamentos comuns entre redes em turismo e outras redes reais. O que é comum a todas as redes na natureza?

Adicionalmente, os padrões encontrados entre os destinos turísticos mostram um comportamento não social, com destinos mais característicos de redes económicas e sistemas tecnológicos que questionam a faceta social do sector do turismo. Por acréscimo, a rede de transportes aéreos e a rede de turismo mostram diferenças consideráveis que se podem dever a razões políticas ou outras que provavelmente explicam o aumento da utilização de voos charters.

keywords

Tourism networks, social network theory, tourism movements growth, graph theory, complex networks.

abstract

In the last decade the concept of networks has been rapidly growing among the tourism literature, generally applied in diverse topics, ranging from inter-organizations, multi-destination structure, tourism webspace, among many others.

The concept of a network is pervasive on nature and society, from economics to management, from biology to medicine, and the knowledge on networks has been developing a common theory for understanding different complex systems, the network representing relations between organizations, stocks, goods, proteins or people.

The thesis aims to have a systematic account of the theoretical achievements on two fertile fields of research that find a common strand now. Complex Networks is a new field on physics that has been strongly developing the last decade, with strong interdisciplinary applications. By other hand, social network analysis has been an active research field in sociology and economics for a long time. The implication of complex networks into the science of tourism networks is a promising area already with fascinating results.

The thesis has three main results. First, it brings knowledge from the rich social and complex network theory. Secondly, it obtains evolution models that better fit the international tourist arrivals. How social networks are organized? How does people decide their travel destinations? These are examples of questions which will be addressed on the thesis.

Third, it discusses results on other tourism and non-tourism real-world networks with common behaviors. What is common to all networks in nature?

Additionally the patterns found between tourist destinations show a nonsocial behaviour of destinations, questioning the social backbones of the tourism sector, and showing similarities with economic and technologic systems. Moreover the air transportation network and the tourism network show dissimilarities that can be taken from political reasons and probably explain the increase use of flights with charters.

This PhD thesis was partially supported by POCI 2010 and FSE.

Ciência.Inovação 2010 Programa Operacional Ciência e Inovação 2010
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Contents

List of Tables

List of Figures

Part I

Introduction

1 Preface

The concept of networks is a fashionable one, whether in nature or society the simple fact that ”everyone is connected to everyone else in life” [Barabasi, 2003], and acting on one another, has some surprising consequences. Castells [1996] states that the rise of the network society driven by the information age defined a new economic age, where the ever-shifting electronic networks, like the international currency market and the Internet, have already emerged as the dominant organizing principle of the new age. One of the most remarkable fingertips of networks is their ability to show relationships among so many seemingly disparate phenomena [Castells, 1996; Barabasi, 2003].

In the last decade the theory of networks pervasive on nature, society, economics, management, biology, medicine, etc, has developed a common theory for understanding different complex systems, being it relations between organizations, stocks, goods, proteins or people. Following this strand the thesis begins with a literature review on the topic of networks on section 3.1, with a historical perspective of graph theory (network theory in mathematics) which started with a simple city planning problem. It follows some of the outstanding results on networks which range from sociology with the theory of the strength of weak ties [Granovetter, 1973] and the six degrees of separation [Travers and Milgram, 1969], to the management theory of structural holes.

An overview is presented on section 3.1.1, from simple evidences of the importance of relations to a general framework of networks, with basic definitions and measurements on section 3.2 and models of the evolution of networks on section 3.3. Network theory recently characterized many real-world networks, like the internet, scientific social collaborations, economic networks, airport networks, among others (see section 9).

Recognizing tourism as a sector strongly influenced by global economic and technologic achievements (see section 2) is understanding its fluidity on the networked society. Among tourism literature, the sector itself is recognized to have strong interdependencies among players, whether they are organizations, tourists, institutions or entrepreneurs, which cooperate and compete in the same space. Generally applied on a diversity of topics, ranging from interorganizations, multidestination structure, tourism webspace, the global framework and theory of networks seems to be widely suitable to understand interdependencies and embeddedness of tourism systems.

A discussion on the contribution of the topic of networks into tourism literature (see section III) results on a general debate on how the different strands emerged into the tourism literature, and defends a general theorization rather than a mere use of methods. The methodological approach to tourism networks is discussed as whether the role of the researcher is to interpret it as a methodology or a theory.

The international tourist arrivals witness one of the fastest growth ever making tourism one of the leading world economic sectors. Competing in the international marketplace is no longer a matter of luck, but rather a strategy based on inner sector knowledge and on society driven forces. Thus, tourism networks is a research field that significantly increased in the last decade, with multiple applications that answer conceptual challenges and industry needs (see section 4).

It is analyzed and mapped the international tourist arrivals/departures network between every two countries in the world [WTO, 2004]. Based on network models the scaling laws of human travel are obtained on section 7, as well as other important characteristics of tourist destination relationships (see section 7.2 and 8), like an empirical evidence of the economic backbones of traveling patterns and methods to analyze the information embeddedness on the network.

A comparison with studies of real-word networks in nature and society (on section 9), and also with other studied networks within the tourism sector, bring very strong evidences of the shape of the travel and tourism sector, as well as its structural similarities and differences with other sectors, providing a quantitative and complementary understanding of networked tourism systems.

List of Publications and Talks

Some of the most fruitful scientific discussions occurred during conferences and while writing publications with researchers from different backgrounds. It follows the list of the publications and talks during the thesis. In the whole list, except number 5, the numerical work when existing, was performed by the author J. I. L. Miguéns. The contribution of the author was also quite significant on literature review, discussion of work choice and results interpretation. A special thank to all the co-authors.

1. J. I. L. Miguéns and J. F. F. Mendes (2008), ”Weighted and Directed Network on Traveling Patterns.” P. Li‘o et al. (Eds.): BIOWIRE 2007, LNCS 5151, 137146.
2. J. I. L. Miguéns and J. F. F. Mendes (2008), ”Weighted Clustering: Finding Roles.” Submitted to the Journal of Statistical Mechanics -Theory and Experiments.
3. J. I. L. Miguéns, R. Baggio and C. Costa (2008), ”Social media and Tourism Destinations: TripAdvisor Case Study.” In: *Proc. Advances in Tourism Research.* Aveiro: University of Aveiro (ISBN ⋯ a acrescentar)
4. J. I. L. Miguéns and J. F. F. Mendes (2008), ”Travel and tourism: Into a complex network.” *Physica A*, 387, 29632971.
5. C. Costa, Z. Breda, R. Costa, and J. I. L. Miguéns (2008), *The benefits of networks to SMEs*, chapter Network Analysis and Tourism: From Theory to Practice, 96114. Aspects of Tourism Book Series, Channel View Publications.
6. J. I. L. Miguéns, A. Corfu (2008), *e-Destination Structure: A Network Analysis Approach*, In: O’Connor/Höpken/Gretzel (eds.), Information and Communication Technologies in Tourism 2008, SpringerWienNewYork. (ISBN 978-3-211-77279-9).

7. J. I. L. Miguéns, J. F. F. Mendes and C. M. M. Costa (2007), ”International tourism network.” In *Proc. SPIE Int. Soc. Opt. Eng.* (SPIE, ed.), 6601.

8. C. Costa, R. Costa, J. I. L. Miguéns (2005). As redes e os clusters como factores de crescimento e internacionalização da economia Portuguesa Competitividade e Sustentabilidade dos Destinos Turísticos, 2ª Conferência do Atlântico. Universidade de Aveiro, Universidade da Madeira e Associação Insular de Geografia, Madeira, $23-27$ de Novembro.

9. J. I. L. Miguéns and C. Costa (2005) *Tourism Networks: new methodologies, potential approach to tourism analysis*, G. Papageorgiou (ed.). Cutting Edge Research in Tourism: New Directions, Challenges and Applications. Guildford: University of Surrey (ISBN 1-84469-012-1).

Evolution of World Tourist Arrivals 2

The evolution of tourism, on business, research or tourist arrivals (clients) is dramatically changing tourism on all its dimensions. On this section an overview of historical developments of tourism, as well as trends drive by scientific developments are presented, reflecting how tourism is approached over the decades emphasizing the relation between tourism, the world economy and scientific achievements.

The substantial growth of tourism activity clearly marks tourism as one of the most remarkable economic and social phenomena of the past century. It has been, in the last decades, one of the economic activities with greater dynamic growth [WTO, 2004]. Tourism faces many research challenges to growth on a sustainable way. The number of international arrivals shows an evolution from a mere 25 million international arrivals in 1950 to over 700 million in 2002, corresponding to an average annual growth rate of 6.6%, see Fig. 2.1. In this way tourism represents approximately 7% of the worldwide exports of goods and services, and is one of the most important sectors of the world's service economy.

These facts reinforce the importance of tourism research [Cooper et al., 2008]. Tourism strongly influences the economy, welfare of population, preservation of cultures, etc. Therefore countries aim to maintain or superpass tourist market shares. Having a knowledge based strategy is fundamental. In 2006, revenues from global tourism were approximately US 733 billion (see Tab. 2.1), resulting from 846 million tourist arrivals [WTO, 2007], see Tab. 2.1. It is expected that

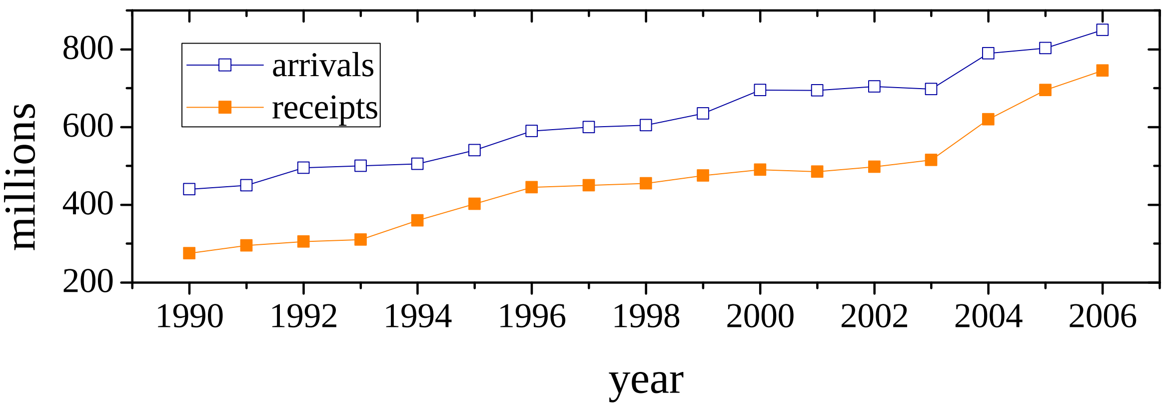


Fig. 2.1: Inbound tourist arrivals (unfilled square) and receipts (full square), 1990-2006. Source: UNWTO.

the size of the global tourism market will reach $1,600$ million people which is equivalent to 20 per cent of the world population by the year 2020 [WTO, 2004]. But, which are the determinants for a country be more attractive than other? What makes it more competitive? How is this huge web growing?

The deployment of tourism, as an economic growth tool requires a comprehensive study, and places tourism as one of the defining phenomena of our age. It has a considerable, and growing, impact on a large range of issues - such as the climate changes and environment, leisure and transport - and in fewer than three hundred years tourism has come to be a global service industry of great economic, cultural and political importance. With tourism sector established as one of the main global economic drivers, its relevance to contribute on the challenges of sustainable development and the response to climate change has also become clear [Smith, 1993].

Tourism has a significant importance for many countries, due to the large intake of money for businesses (see Tab. 2.1) with their goods and services and the opportunity for employment in the service industries associated with tourism. The sector itself can be subdivided in service industries including an overarching business comprising hundreds of component businesses, some huge but mostly small businesses, including airlines, cruise lines, railroads, rental car agencies,

Tab. 2.1: International Tourism Receipts. Source: UNWTO.

Rank	billion		change (%)		Local Currency Change	
	2005	2006	05/04	06/05	05/04	06/05
United States	81.8	85.7	9.7	4.8	9.7	4.8
Spain	48.0	51.1	6.0	6.6	6.0	5.6
France	42.3	42.9	3.5	1.5	3.5	0.6
Italy	35.4	38.1	-0.7	7.7	-0.7	6.7
China	29.3	33.9	13.8	15.9	13.8	15.9

travel marketers and expenditures, lodging, restaurants, convention centers, travel reception services, commercial campground, and parts of retail shops, food stores, and gas stations.

Tourism as a concept can be viewed from different perspectives. It is an activity in which people are engaged in travel away from home primarily for business or pleasure. It is a business providing goods and services for travelers, and involves any expenditure incurred by or for a visitor for his or her trip. Therefore, tourism is an umbrella of concepts, and as old as tourism takes place, tourism has been changing its definition. The difficulty on developing a definitions is due to the several dimensions on a tourism system, where the complexity of interactions and consequences that occur before, during and after traveling, with psychologic, sociological, ecological, and political impacts (see section 5.2).

Tourism is a science with a full complexity of interactions, a long body of knowledge and a constituency of millions of tourists who feel themselves a part of the tourism institution. For all this complexity of interactions, defining tourism is subjective, as taking the point of view of a researcher, tourist, government body, hotel manager, transportation manager, etc, considerable changes the goal of the author, affecting the definition. One of the first definitions of tourism goes back to the XIX century, as ”people that travel for the pleasure of traveling, out of curiosity, and because they have nothing better to do, for the joy of boasting about it afterwards” [Sigaux, 1876].

The first definitions of tourism were related with the demand side of the sector, the way tourist experience and stay out of their usual environment. Hunziker and Krapf [1942] defined

Tab. 2.2: World's Top Tourist Arrivals, where Pop stands for population. Source: UNWTO.

					Market %		**Arrivals per**
Rank	**2003**	**2004**	**03/02**	**04/03**	**2004**	**Pop.**	100 **of Pop.**
World	**693**	**764**	**-1.9**	**10.2**	**100**	**6,377**	**11**
France	75.0	75.1	-2.6	0.1	9.8	60	124
Spain	50.9	52.4	-2.8	3.1	6.9	40	138
Unite States	41.2	46.1	-5.4	11.8	6.0	293	17
China	33.0	41.8	-10.4	26.7	5.5	1,299	4
Italy	39.6	37.1	-0.5	-6.4	4.9	58	64
United Kingdom	33.0	41.8	-10.4	26.7	5.5	1,299	4
Mexico	18.7	20.6	-5.1	10.5	2.7	105	21

tourism as "the sum of the phenomena and relationships arising from the travel and stay of non-residents, insofar as they do not lead to permanent residence and are not connected with any earning activity" [Hunziker and Krapf, 1942]. In 1976 Tourism Society of England defined it as "Tourism is the temporary, short-term movement of people to destination outside the places where they normally live and work and their activities during the stay at each destination. It includes movements for all purposes."

Over the years, the classifications and methodologies on tourism develop to fit industry impact, and the supply side starts playing a role on tourism concepts, as a common quantitative way was needed to measure the economic impact of tourism. In 1994 the United Nations classified three forms of tourism in its Recommendations on Tourism Statistics: Domestic tourism, which involves residents of the given country traveling only within this country; Inbound tourism, involving non-residents traveling in the given country; and Outbound tourism, involving residents traveling in another country. The United Nations also derived different categories of tourism by combining the three basic forms of tourism: Internal tourism, which comprises domestic tourism and inbound tourism; National tourism, which comprises domestic tourism and outbound tourism;

and International tourism, which consists of inbound tourism and outbound tourism. On section 6 definitions from the UNWTO of tourism are introduced.

The history of European tourism can perhaps be said to originate with the medieval pilgrimage. During the 17^{th} century, it became fashionable in England to undertake a Grand Tour. Mass travel could only develop with improvements in technology allowed the transport of large number of people in a short space of time to places of leisure interest, and greater numbers of people began to enjoy the benefits of leisure time. The technologic developments brought a new railway network, and the spread of railway network in the 19^{th} century influenced the growth of Britain's seaside towns. Increasing speed on railways meant that the tourist industry could develop internationally. To this may be added the development of sea travel. The relation between the air transportation and the tourism industry are discussed on section 9.2. The age when tourism reached a significant number of international mass travel began with the growth of air travel after World War II.

Along with some low sustainable developed tourism destinations also other setbacks affect tourism, like the terrorism attack on September $11, 2001$ and following threats on tourist destinations of Bali and European cities. The growing competitiveness of tourist destination require more and more a continually improvement and adaptation. Also natural disasters can dramatically destroy a destination, such as the tsunami, caused by the 2004 Indian Ocean earthquake that hit Asian countries bordering the Indian Ocean. Along with the lack of sustainable planning and possible natural disasters, tourism also faces the threat of global warming. The impacts of travel and tourism sector are responsible for 5.3% of global CO_2 emissions, [Gössling, 2002; Smith, 1993]. Among tourism related activities transportation contributes with 94% of CO2 emissions, placing it among the most contributor to global warming. It is important noticing that the Kyoto protocol [Nations, 1998] does not cover aviation sector, which responds for 40% of the emission of CO_2 on travel and tourism transportation. An indicator of how natural disasters, weak economy and wars are affecting tourism industry is proposed on section 8.3.

The continuously challenges brought by the surrounding environment of the tourist and the industry keeps with huge trends [Buhalis and Costa, 2005], that did, do and will modify tourism. From the mass tourism to package tourism, the amalgam of new forms of tourism keeps on growing with the world GDP significant increase, and emergent tourism destination that star playing a central role, such as the most significant case of China Tab. 2.1 and Tab. 2.2. Understanding the evolution of the world tourist arrivals is one of the main goals of this thesis, that can bring more informed decisions of strategic position of countries as international tourist destinations. The air transportation sector have made tourism more affordable, as low-cost airlines keep on lowering prices (relation between charters and tourism discussed on section 9.2).

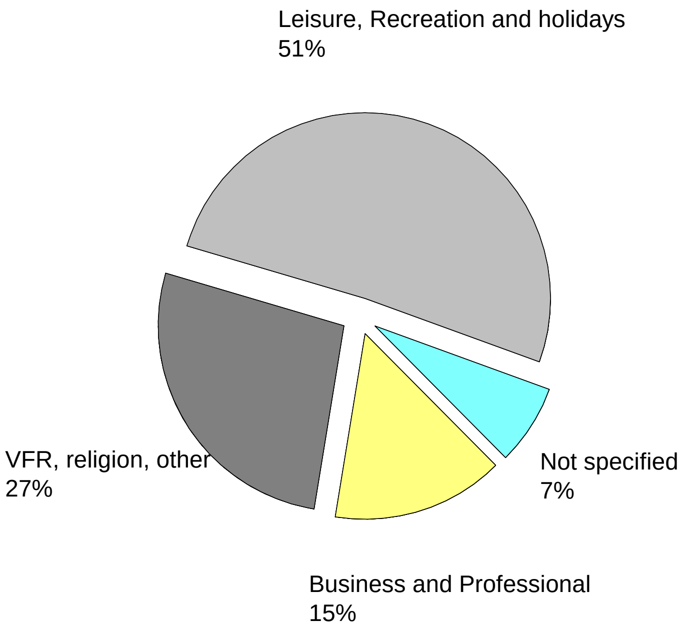


Fig. 2.2: Inbound tourism by purpose of visit, 2006. Source: UNWTO.

The United Nations World Tourism Organization (UNWTO) forecasts that international tourism will continue growing at the average annual rate of 4% [WTO, 2004]. By 2020 Europe will remain the most popular destination, but its share will drop from 60% in 1995 to 46%. Long-haul will grow slightly faster than intraregional travel and by 2020 its share will increase from 18% in 1995 to 24%. The year of 2007 proved the resilience and potential of international tourism and 2008 looks likely to confirm the solid development of the sector [WTO, 2007]. With 846 million

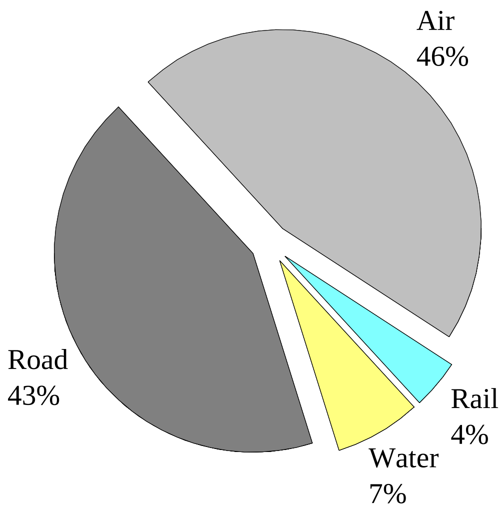


Fig. 2.3: Inbound tourism by means of transport, 2006. Source: UNWTO.

international tourist arrivals, corresponding to an increase of 5.4% over the previous year, 2006 exceeded expectations. The tourism sector continued to enjoy above average results and recorded a third year of sustained growth. One notable feature of 2006 was the continuing healthy performance of emerging destinations, backed up by one of the longest periods of sustained economic expansion. All regions and subregions succeeded in achieving positive growth, although the regional averages mask some fairly mixed performances across different subregions and countries.

In 2006, just over half of all international tourist arrivals were motivated by leisure, recreation and holidays (51%) a total of 430 million. Business travel accounted for some 16% (131 million), and 27% represented travel for other purposes, such as visiting friends and relatives, religious reasons/pilgrimages, health treatment, etc (225 million). Air transport (46%) and transport over land whether by road (43%) or rail (4%) generate roughly equal shares of all arrivals, while arrivals over water accounted for 7% in 2006, see Fig. 2.3. For the past three years, the trend has been for air transport to grow at a faster pace than ground and water transport.

Experience shows that in the short term, periods of faster growth (1995, 1996, 2000) alternate with periods of slow growth (2001 to 2003). While the pace of growth till 2000 actually exceeded the Tourism 2020 Vision forecast [WTO, 2007], it is generally expected that the current slowdown will be compensated in the medium to long term. The analysis of the world tourism network has patterns, that are measured, depicting the slow down periods of tourist arrivals (see section 8.3). Therefore, a proposed measurement for governments to monitor the changes of their competitive position and slow down effects is presented.

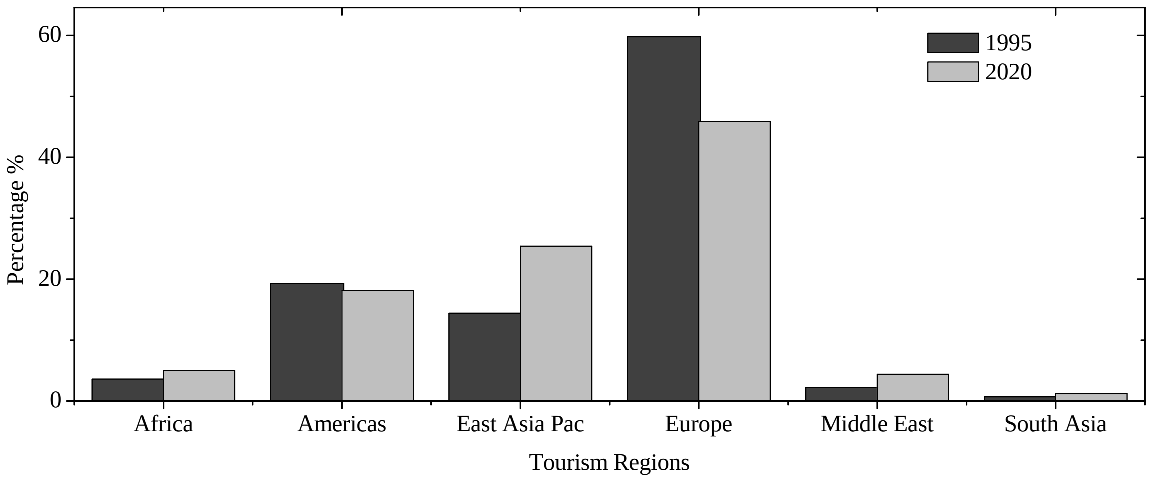


Fig. 2.4: Forecast of tourism regions share in 2020. Source: UNWTO.

The World Tourism Organization forecasts that among the 1.6 billion tourists expected by the year 2020, 1.2 billion will be intra-regional and 378 million will be long-haul travelers. The total tourist arrivals by region shows that by 2020 the top three receiving regions will be Europe (717 million tourists), East Asia and the Pacific (397 million) and the Americas (282 million), followed by Africa, the Middle East and South Asia, see Fig. 2.4.

East Asia and the Pacific, Asia, the Middle East and Africa are forecasted to record growth at rates of over 5% year, compared to the world average of 4.1%. The more mature regions Europe and Americas are anticipated to show lower than average growth rates. Europe will maintain the highest share of world arrivals, although there will be a decline from 60 per cent in 1995 to 46 per cent in 2020.

2.1 Online Social Nets on Tourism

The access to tourism products, mainly flights and hotels is strongly influenced by Information, Communication and technologic (ICT) trends by facilitating the purchasing of tourism products, through internet and in the next years through mobile, transforming the space-time relation between the tourist and the destination. With the advent of e-commerce, tourism products have become one of the most traded items on the internet. Tourism products and services have been made available through intermediaries, although tourism providers (hotels, airlines, etc.) can sell their services directly. This has put pressure on intermediaries from both on-line and traditional shops [Buhalis and Main, 1998]. A new form of tourism is taking place with technologic improvements. Space tourism already an expectation, is about to start in the first decades of the 21^{st} century. The space as a destination brings the expectations also on air-ship hotels.

Some tendencies reveal that we will have computerized booking systems [Holder, 1991], recommender systems for traveling planning [Fesenmaier et al., 2006], online communities [Kim et al., 2004; Wang and Fesenmaier, 2002], virtual tours [Buhalis, 1998], web-based interpretation to encourage visitation [Bèdard et al., 2008], etc. Information and communication technologies can strongly influence the way tourism is experienced and destinations selected [Buhalis and Law, 2008; Buhalis, 1998; Sheldon, 1998; Poon, 1993].

Tourism informatics (or eTourism) is a multi-billion dollar international industry. It is also one of the biggest users of web technologies and constantly adopts innovative ideas to enhance its market penetration. The World Wide Web is currently undergoing a further revolution. While e-commerce played a key factor at the end of $1990s$, a new form of collaborative activity emerges online today. Rather than more or less sophisticated e-commerce platforms, Web 2.0 business models provide services that invite users to a direct and strong participation and derive profitable returns from the several forms of advertising present online [Fogelman-Soulie and Herault, 1989]. The ever growing rate of Internet diffusion is still happening at a fast pace, so that the new forms of online social networking are unsure yet of their future, and there is an on-going discussion about the consequences and the effects of social network sites, both for practitioners and researchers [Loudhouse-Research, 2007].

Tourism is a sector with a close relationship with the new information and communication technologies. It is deemed that a good understanding of the quality and quantity of the mechanisms for spreading information online can facilitate tourism managers (whether of a Destination-Management-Organization an hotel or any other tourism related company) to market effectively own organization online. An important feature of Web 2.0 applications is the rich wealth of user generated content. This can prove highly influential in directing tourists' choices, but can be also of extreme value for the comprehension of preferences, needs and reactions which can (or should) inform many decisions from a management point of view.

Social networks are online communities of people who share common interests and activities. They provide a user with a collection of various interaction possibilities, ranging from a simple chat, to multiple video conferences, and from the exchange of plain email messages to the participation in blogs and discussion groups. Online social networks may also contain categorized relationships (e.g. former classmates), means to connect with friends (with self assembled description pages), or recommendation systems for some kind of objects or activities. Popular spaces combine different functions of this type. Some of the most widely attended are systems such as MySpace (190 million users in 2007), Orkut (over 62 million), or LinkedIn (over 5.5 million).

The Web2.0 is strongly characterized by an unprecedented easiness of interactivity which fosters the formation of communities and the generation of user-driven content. Its diffusion has been rapid and widespread, so that today, for example, blogs are counted in billions. It is no surprise then to find out that travel and tourism related topics are among the most popular issues in this environment. Travel plans, destinations and hotels reviews, tourist guides, suggestions for restaurants or exhibitions are ever growing discussion subjects and the term Travel 2.0 has started denoting this trend. Tourism on the Internet was already one of the major 'players' [Buhalis, 1998], and the online travel market has assumed very a consistent size. In Europe, for example, it represents (in 2007) almost 19.4% of the total market [Gretzel et al., 2008]. Moreover a continuing sensible growth is predicted for the next years.

The World Wide Web is currently undergoing a further revolution. While e-commerce played a key factor at the end of 1990s, a new form of collaborative activity emerges online today.

Rather than more or less sophisticated e-commerce platforms, Web 2.0 business models provide services that invite users to a direct and strong participation and derive profitable returns from the several forms of advertising present online [Fogelman-Soulie and Herault, 1989]. The Internet diffusion and technological evolution is still happening at a fast pace, and the new forms of online social networking are unsure yet of their future. There is an on-going discussion about the consequences and the effects of social network sites, both for practitioners and researchers [Loudhouse-Research, 2007].

Tourism is one of the sectors with a very close relationship with the new information and communication technologies. It is deemed that a good understanding of the quality and quantity of the mechanisms for spreading information online can facilitate tourism managers (whether of a DMO an hotel or any other tourism related company) to market effectively own organization online [Buhalis, 1998].

An important feature of Web 2.0 applications is the rich wealth of user generated content. This can prove highly influential in directing tourists' choices, but can be also of extreme value for the comprehension of preferences, needs and reactions which can (or should) inform many decisions from a management point of view.

Web 2.0 applications in the tourism sector have been named Travel 2.0 applications by Philip C. Wolf (president and CEO of PhoCusWright, a leading consultancy firm in the tourism arena) and are creating, for the umpteenth time, a cultural change in the tourism world. The traditional operators are facing a new consumer, which can easily access information and easily share their views, comments and suggestions in an informal and collaborative way, increasing the value and influential power as determinants of choice for other consumers. The Web is shifting to a business-to-consumer marketing to a peer-to-peer model for the sharing of information. All tourism businesses are thus facing the need to implement strategies and tools (websites or portals) based on user generated contents or, at least, to incorporate these new technologies to enrich their multimedia contents.

Examples are already available online. The Tourism British Council has been one of the first destinations to include blogs and user generated content in their marketing strategy, and the

Florence Official Tourist Office (like many others) embraces geo-referenced contents regarding about tourism attractions through GoogleMaps [Firenze-Turismo, 2008].

Online social travel networking is also changing the way tourists plan their trips. These website allow users to interact and, for example, provide reviews on hotels or on local tourist attractions. Some examples of these websites are TravBuddy.com, Travellerspoint, WAYN, Woophy, Passportstamp, and TripAdvisor.com. The latter is probably the largest travel community on the Web. It was founded in 2000 and currently it covers 212000 hotels, over 30000 destinations, and 74000 attractions worldwide [Sheet, 2008].

Social network analysis investigates such interactions between people, groups and organizations [Watts, 2003a; Wasserman et al., 1994; Scott, 2000]. By disseminating information via their social networks, individuals create strong peer influence that often surpasses exogenous influences. Marketing leverages this peer influence to trigger self-reinforcing content propagation among individuals [Litvin et al., 2008]. Weak and strong ties [Granovetter, 1973] between those individuals determine the distinct paths of information dissemination. Travel and communication networks are essential to facilitating participation in social networks and for the generation of social capital [Larsen et al., 2006]. The need to spend time traveling is an inevitable consequence of co-present (face-to-face) obligations that are embedded in social practice (e.g. business, social or legal events) [Litvin et al., 2008; Larsen et al., 2006].

Increasingly, ICTs enable travelers to access reliable and accurate information as well as to undertake reservations in a fraction of time, cost and inconvenience required by conventional methods [O'Connor, 1999]. ICTs can assist in the improvement of the service quality and contribute to higher guest/traveller satisfaction. ICTs place users in the middle of its functionality and product delivery. Every tourist is different, carrying a unique blend of experiences, motivations, and desires. To an extent the new sophisticated traveler has emerged as a result of experience. Tourists from the major generating regions of the world have become frequent travelers, are linguistically and technologically skilled and can function in multi-cultural and demanding environments overseas.

2.2 Socioeconomic Trends

There is increasing competition in the tourism industry, whether we consider international destinations (section 7), or domestic destinations, but also the among firms within (in between) destination(s). The knowledge plays a central factor to develop destination as tourism develops the greater the capacity of destination managers and tourism operators to formulate strategies to achieve competitive advantage for their organizations.

The major shifts in the tourism industry are shaped by society, scientific and technologic changes, which affects consumer behavior and organizations like also the political forces, environmental shifts and strongly the great growth of information and communication technology. The organizations face the need to a 'strategic drift' [Buhalis, 1998; Johnson and Scholes, 1997].

The world driving forces are strongly impacted by globalization, where easier access across borders, affects tourism industry, with an increase of foreign tourists as well as increased global competition from international tourist destinations. This thesis quantifies the evolution growth of the world net of tourists, between every two countries in the world. The world economy is forecasted to grow over significantly the next decade and a half [Third-World-Network, 2006]. While the projected dynamic world economy is forecast to provide the basis for increased international and domestic tourism, the growing dynamics net relies on a more precise quantification of this evolution.

One of the strongest driving forces in tourism is the rising of income, remarkably generating tourism flows [Crouch, 1994]. Also the demographic shifts and social changes are having profound effects on almost every social institution [Dwyer and Kim, 2003]. The demographic is witness the aging of population, more adults want to be teenagers, breakdown of traditional family grouping, more single parents, people marrying on a later age, all affecting the way people take holidays [Hall, 2000]. The level of qualification of tourists is related with the requirements of globalizing economy and technological change inevitably require a more highly skilled labor force. Destinations and organizations should increase education as a determinative of success. This includes innovative businesses that are well attuned to their customers needs and staffed with highly educated workers valued as 'human capital' and organizations with external knowledge focusing on organizational culture that enshrines life-long learning [James, 1997].

Along with demographic changes there are also shifts on aspirations and expectations that affect values of consumers in diverse ways. In general in developed world populations is more individualistic, willing to have a new experience, educated, higher demand for short time holidays, seeking for price quality relation, environmentally aware, and safety conscious [Dwyer and Kim, 2003; Willmott and Graham, 2001]. One result of the experience economy and tourism has been a fragmentation of the tourist market into subsets of unique experiences [Elliot and Johns, 1993].

Increasingly, tourists are demanding assurances of safe products and services prior to purchase. Since tourist behavior is as constrained by perceived risk as it is by actual risk operators need to address perceptions of risks as well as the risks themselves [Lepp and Gibson, 2003].

Another worldwide trend is urban congestion , both in the industrialized and developing worlds tourism, and travelers will tend to favor holidays away from crowds, increasingly the need to engage in discretionary tourism to escape and/or to indulge [WTO, 2007]. Cities will need to work hard to develop in a way that to make them destinations that are worth visiting for more than a short break, and for more than one time visit. Another trend is the high standard of public health in developed countries, contributing to increase longevity, influencing demand for a combination of health and travel products. In developed countries there will be a blurring of working life and retirement [Cetron, 2001].

Part II

Networks: From Practice to Theory

3 Network Breakthroughs and Complex Networks

Look with all your eyes, look
(Jules Verne, *Michael Strogoff*)

Introduction

Without intending to review in detail the full history of network theory, an overview of the main developments of the areas that most contributed to network theory are presented in this section. These theories and methodologies bring new tools promising to the travel and tourism industry.

The concept of a network lies on the way a system is perceived, whether we deal with the linking between molecules or people, a natural or economic system – is not a sum of elements to be distinguished from other and analyzed individually discretely, but a pattern, or a structure: the element's existence does not precede the existence of the whole, where the simple parts do not determine the pattern, but the pattern determines the parts. We could compare it to a puzzle, with specific coloring and shape, where all the pieces are observed. The pieces are readable only when assembled and in isolation a puzzle piece means nothing. But when fitting the piece into one of its neighbors, the piece as an individual disappears, ceases to exist as a piece, becoming a part of a whole.

Historically, the study of networks has been mainly the domain of a branch of discrete mathematics known as graph theory. Since its birth in 1736, when the Swiss mathematician Leonhard Euler published the solution to the Königsberg bridge problem. The city problem was to find a round trip that crossed each bridge of Königsberg once and only once). Since then graph theory has witnessed many key developments and has provided a framework to answer several practical questions like, calculating maximum flows per unit time from source to sink in a network of pipes, how to color regions of a map using the minimum number of colors and considering that neighbor regions receive do not have the same color, or how to fill n jobs by n people with maximum total utility.

In general a graph is a representation of a set of nodes wired by edges, forming a net. Later some researchers had worked to understand which were the properties of some types of graphs and its process of construction, that is, aiming to model the process of grouping [Albert and Barabási, 2002; Watts, 2003b; Buchanan, 2003]. For example we as individuals are the units of a network of social relationships of different kinds. With the developments in mathematical graph theory, the study of networks has seen important achievements in some specialized contexts, as for instance in the social sciences, named social network analysis (SNA) [Wasserman et al., 1994; Scott, 2000], with a common toolbox for describing network structures borrowed from SNA [Pforr, 2006].

Social networks analysis started to develop in the early 1920s and focuses on relationships among social entities, as communication between members of a group, trades among nations, or economic transactions between corporations. Moreover the last decade has witnessed the birth of a new movement of interest and research in the study of complex networks, with some complex structure and dynamically evolving in time, with the main focus moving from the analysis of small networks to that of systems with thousands or millions of nodes. In this sense the attention moved to the properties of networks of dynamical units. Hence networks in a more general context can represent electric power grids, the Internet, highways or subway systems, and neural networks, etc. Or they can be entities defined in an abstract space, such as networks of acquaintances or collaborations between individuals.

The network theory was triggered by two main works, the first by Watts and Strogatz on small-world networks [Strogatz, 2001] in Nature on 1998, and the other by Albert and Barabási [1999] on scale-free networks appeared in 1999 in Science, the flourish amount of applications has been certainly induced by the increased computing powers and by the possibility to study the properties of large databases of real networks. These include transportation networks [Guimera et al., 2005], phone call networks, the Internet and the World Wide Web [Albert et al., 1999], the actors' collaboration network in movie databases, scientific coauthorship [Newman et al., 2002], and also systems of interest in biology and medicine, as neural networks, metabolic and protein networks [Wuchty and Almaas, 2005].

The massive and comparative analysis of networks from different fields has produced an amount of unexpected interesting results. The first issue that has been faced is certainly structural. The research on complex networks began with the effort of defining new concepts and measures to characterize the topology of real networks. The main result has been the identification of a series of unifying principles common to most of the real networks considered, which we shall present on the following sections.

A relevant property regards the degree of a node, that is the number of its direct connections to other nodes. In real networks, the degree distribution $P(k)$, defined as the probability that a node chosen at random has degree k or, equivalently, as the fraction of nodes in the graph having degree k, significantly deviates from the Poisson distribution expected for a random graph and, in many cases, exhibits a power-law (scale-free) tail. Moreover, real networks are characterized by correlations in the node degrees, by having relatively short paths between any two nodes (small-world property), and also by the presence of a large number of short cycles.

These empirical findings have initiated a new era of network modeling, since the models proposed initially in graph theory turned out to be quite far from the real-world networks characterization. Scientists had to do with the development of new models the growth of a network to better fit the structural properties observed in real-world networks topologies. The structure of a real network is the result of the continuous evolution or dynamics of the forces that drive it, and definitely affects the function of the system. So that starting stage of the reserach on complex networks was motivated by the goal of understanding and modeling the structure of real-world network and that it would lead to a better knowledge of its evolutionary mechanisms, and to a better understanding on its dynamical and functional behavior.

3.1 Graph Theory: Historical Overview

Graph theory was designed to test the ingenuity and solve local problems rather than usually in mathematics that problems are stimulated by the capacity of abstraction. But despite the apparent triviality of such puzzles, they captured the interest of mathematicians, with the result that graph theory has become a subject rich in theoretical results of a surprising variety and depth. The origin of graph theory has its ramifications on one particular problem - Königsberg bridges. The solution of this problem involves the formulation of several of the basic concepts of graph theory.

The Seven Bridges of Königsberg is a problem inspired by an actual city and citizens situation. The city of Königsberg (presently Kaliningrad, Russia) is crossed by the river Pregel, and includes two islands which were connected to each other and the mainland by seven bridges (see figure 3.1). The problem is whether it is possible to walk with a route that crosses each bridge exactly once, and return to the starting point. It is said that prosperous and educated townspeople allegedly walked about on Sundays trying to solve the problem, but this might be an urban legend. In 1736, one of the leading mathematicians of the time, Leonhard Euler wrote an article in which he dealt with this particular problem and gave a general method for other problems of the same type. His article was of considered importance, both for graph theory and for the development of mathematics as a whole.

Euler's treatment of the Königsberg problem involved two major steps. First he replaced the map of the city (see Fig. 3.1) by a simple diagram showing its main features (see Fig. 3.2), and then, he formulated the problem in such a way that the diagram became unnecessary. Nevertheless, the pictorial representation of graphs is a very useful technique (see Fig. 3.2). He denoted the four land areas by the symbols A, B, C, D and the seven bridges by a, b, c, d, e, f, g, where the bridges a joins A and B, f joins A and D, and so on. This is an example of what we now refer to as a graph,

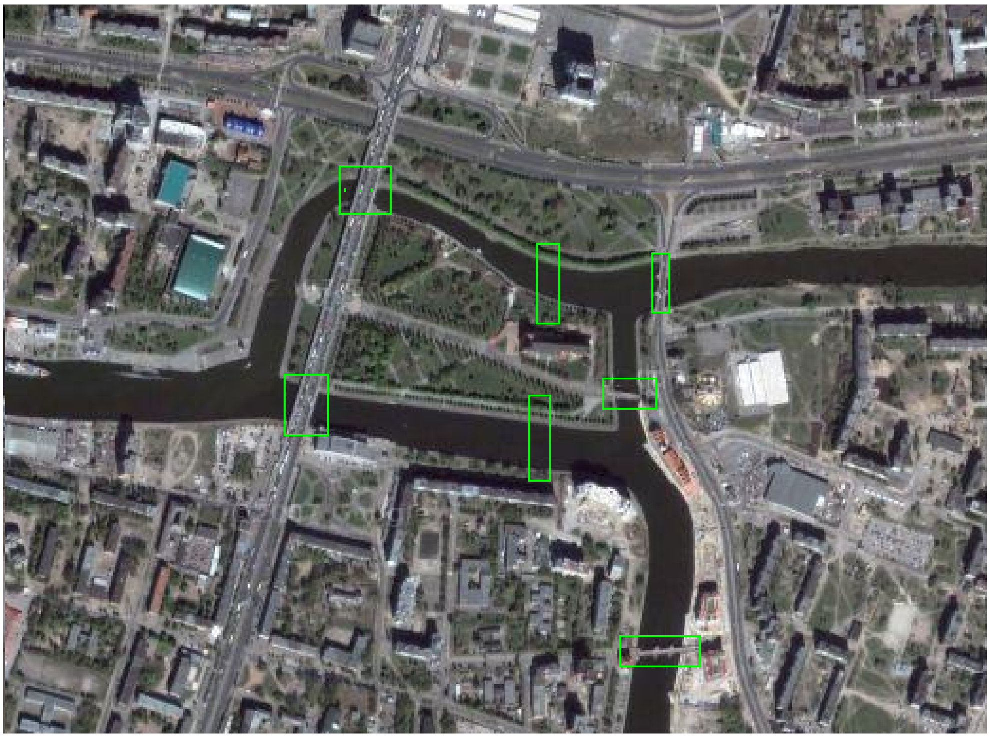
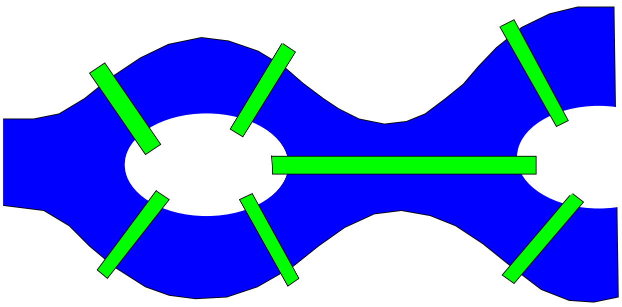
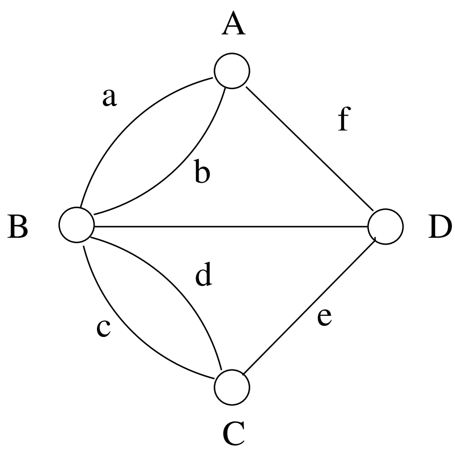


Fig. 3.1: Königsberg bridge problem. Image from Google Maps on Kaliningrad, Russia.



(a) Pictorial representation of the problem.



(b) The graph: nodes and edges formulation of the Königsberg problem.

Fig. 3.2: The graph representation of Königsberg bridge problem, from pictorial image.

and Euler's problem of finding a sequence of eight symbols with a particular property is related to the existence of a special kind of path in the graph.

To explain the meaning of these terms some definitions are introduced. A network (or graph) consists of a finite set of nodes, edges, and a rule which tells the edges connecting to which pairs of nodes. In our particular example there are four nodes, corresponding to the land areas A, B, C, D, and seven bridges. The rule tells them that the edges a and b join the vertices A and B, the edges c and d join the vertices A and C, and so on.

It is helpful to illustrate these abstract definition by representing a graph pictorially. We depict a graph as a diagram of points and lines, in which the points represent vertices and the lines represent edges; a diagram for the Königsberg graph is shown in Fig. 3.1. It should be noted that this is merely a convenient way of describing the graph – we repeat that the graph itself is an abstract entity consisting of the four vertices A, B, C and D, the seven edges a, b, c, d, e, f, g, and the rule which tells us how the edges join the vertices 3.2.

The Königsberg bridge problem is also regarded as one of the first topological results in geometry; that is, it does not depend on any measurements. This illustrates the deep connection between graph theory and topology. Afterwards graph theory the investigation of graphs arose in a more mathematical way, from the study of operators and differential calculus and the interchange of ideas between different branches of science was often highly beneficial to all of them. Graph theory had also a great development through the collaboration between chemistry and mathematics.

3.1.1 Complex Networks: Motivations

Recently the research field of networks has been growing on attention due to the complex networks through studies of the applied areas of the physics to the social networks [Watts, 2003a,b; Albert and Barabási, 2002; Newman, 2000; Amaral et al., 2000] and to networks as a whole [Barabasi, 2003; Dorogovtsev and Mendes, 2003; Buchanan, 2003; Watts, 2003a]. In these new perspectives in an attempt to explain characteristics and properties of real world networks is growing of new field of knowledge.

One of the first studies on graphs and its properties was made by Paul Erdös and Alfred Rényi [Watts, 2003b; Albert and Barabási, 2002; Buchanan, 2003]. They had develop the theory of networks, amongst which the theorization of ”random graphs”. Erdös and Rényi had still attempted against for another fact: the more links was added, bigger the probability to be generated clusters, that is, more hardwired groups. A party, therefore, could be a set of clusters (groups of people) that from time to time they established relations with other groups (networks). It was believed that the process of formation of the graphs was random, the connection were made on a random choice. Erdös and Rényi have defended that the all of us on a social network would have around the same amount of connections, or equal possibilities to receive new links (model description on section 3.3.1). On their point of view, the bigger the network the bigger its fit to a random network.

By observing social networks as a special type of networks it was believe that all people would be linked to the others in the some level, or with the same amount of connections. Sociologist Stanley Milgram [Travers and Milgram, 1969], in the sixties, was the first one to carry through an experiment to observe the degrees of separation between people [Degenne and Forsé, 1999; Buchanan, 2003; Albert and Barabási, 2002; Watts, 2003b]. He sent a certain amount of letters to randomly chosen individuals, and asked them to send it to a specific target. In case that they did not know the target, the chosen people were requested to send the letters for the person they believed to be more close to the final target. Milgram discovered that, by analyzing the letters that arrived to its final address, that the majority had passed only for a small number of people. The result was quite surprising. Most of the letters went through a very small number of people, when it was considered people from distant counties, living thousands of kilometers away, having no clue on the other party. In conclusion, Milgram discovered that we are all just a few degrees of separation from each others, that is, in one ”small-world”. This result is also known as ”six degrees of separation”, the few steps that in average separate every two person in the country.

Another important contribution for the problem of the structure of the social networks was given by the sociologist Granovetter [1973]. In his study, he discovered how important are weak ties or that they posses the most important information on a social structure, when compared to strong ties, for which sociologists were giving more importance.

Granovetter also showed that people who shared strong ties in general (with friends, for example) participated of one same social cluster (of one same group that highly would be clustered). The people within a group that have weaker ties are significantly important for information and power flows because they connect different social groups. Without them, the several clusters would exist as isolated islands and not as a network. The work of Granovetter again reinforces the importance of the triads in the social networks. The social networks, therefore, are not simply random. Some type of order in them exists.

From the experiment of Milgram and the theories of Granovetter, Duncan Watts and his supervisor, Steven Strogatz [Watts, 2003b], had discovered that the social networks they presented hardwired standards highly, tending to form small amounts of connections between each individual (see section 3.3.3). They had created a similar model to the one of Erdös and Rényi, where the ties were established between people next by, and some others established ties in random way. In this way the network was transformed into a small-world [Watts and Strogatz, 1998; Watts, 2003b].

This model would demonstrate that the distance average for any two individuals would not exceed a small number of connections to other individuals, being enough that some random ties between groups still exist [Buchanan, 2003]. The model of Watts and Strogatz shows a network model of real social networks: each one of us has close by friends and also friends that known in some places of the world, that in turn, have other known friends. On a large scale, these last connections show the existence of few degrees of separation between all individuals. Moreover, they had shown that few connections were enough to enters several different clusters to form a small-world in a large network, converting the small networks into a large one [Buchanan, 2003].

The first problem of the theory of the small-worlds of Watts was demonstrated by Albert and Barabási [1999] shortly after the publication of the work: Watts treated its social networks as random networks, that is, networks where the connections between individuals were established in random way, accurately as Erdös and Rényi years before. However, Albert and Barabási [2002] demonstrated that networks were not formed in a random way. He believed that, as the studies of Watts and Strogatz, as well as of Granovetter had pointed, there existed an order in the dynamics of the networks with some well specific laws.

This law, or structural pattern, was called by Barabási ”preferential attachment”. That is, the more connections a node possess, higher the probabilities of having new connections, therefore a new node tends to connect with the more connected ones. This also implies that the nodes would not have the same number of connections. In contrast, such networks would possess few with highly connected nodes (hubs or connectors) and a great majority of the nodes with few connections. Hubs would be ”more popular” and tend to always receive more connections. The networks with these characteristics has been called ”scale-free” [Albert and Barabási, 1999].

The model of Barabási and Albert, for example, has an average degree with low connectivity, since only some nodes are highly connected, the majority has few links. The analytic solution of this model was obtained by [Dorogovtsev et al., 2000]. Moreover, a scale-free network is not necessarily a small-world. However the model of Watts and Strogatz has a connectivity degree similar to the one of a random graph [Erdös and Rényi, 1960], but has high degree of connection for the nodes. In the real world, the networks usually show a degree of distribution (connectivity) varied, that are not necessarily fitting the proposed models (see section 3.3).

Most the relations between the individuals in social and nature networks are not random. The people take in account diverse factors when choosing to connect themselves or not to somebody. The model of the Albert and Barabási [1999] brings important insight in the direction of to foresee the mechanism of construction of the networks, of ”preferential attachment” and the presence of connectors.

However, how do these models give account of the phenomenon of networks in the travel and tourism industry and the movement of humans traveling worldwide? Network theory has rapidly moved across different domains of knowledge and expertise over the last decades. On this thesis is investigated some strands in that movement with two goals in mind. One concerns the issue of how we use network theory in tourism research. The other concerns the patterns of a self organized structure in international tourist destination.

3.2 Definition and Classification of a Network

On this section the main definitions of network theory are introduced, it is also referred some specific relation between theories and their use for the data analysis of this thesis on section 6.

3.2.1 What is a Network?

A network is a set of nodes that are related through a set of relationships. More formally, a network contains a set of objects (called nodes) and relations between the nodes, see Fig. 3.3. Nodes might be people, organizations, computer routers, airports, and the relationship that links them might be for example friendship, economic transactions, information exchange.

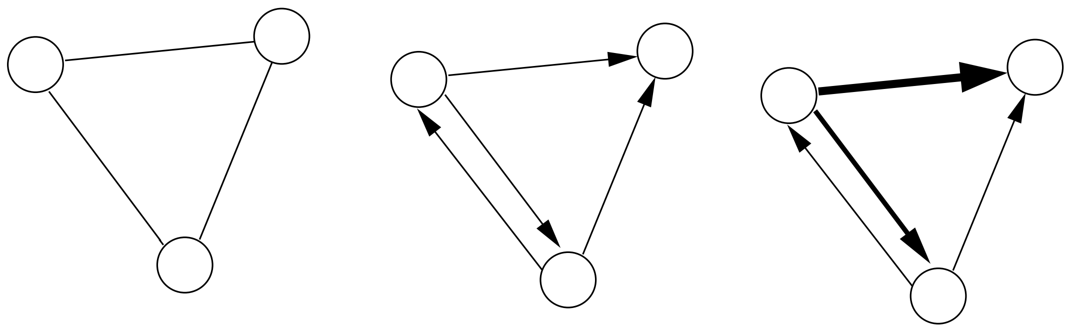


Fig. 3.3: A simple network example. The representation of nodes and relationships, respectively on a undirected and unweighted graph on the left, a directed and unweighted graph on the middle and a weighted and directed graph on the right.

The network can be fully represented by its adjacent matrix and every property of the network can be extracted from its adjacent matrix:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{pmatrix}$$

for N total number of nodes and a_{ij} :

$$\mathrm{a}_{ij} = \begin{cases} 1 & \text{if vertex i is connected to j,} \\ 0 & \text{otherwise.} \end{cases}$$

The first measurement on a network is the degree of a node. Social network researchers measure network importance for a node by the number of connections a node has, named degree. Adjacent matrices in particular, are used to show managers a simple, systematic way to quantify proximity, to lay the foundation for developing a strategic plan. The centrality of nodes on a network is of primary importance, as more competitive nodes have better strategic positions [Burt, 1995; Wasserman et al., 1994; Scott, 2000]. Several measures of centrality have been developed, like degree centrality, closeness, betweenness, eigenvector centrality, information centrality, among others. Centralization refers to the extent to which a network revolves around a single node.

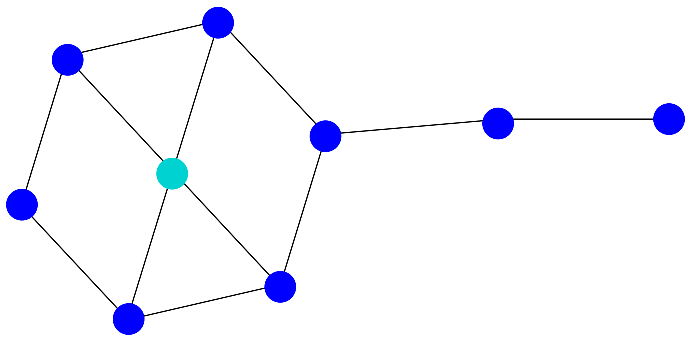


Fig. 3.4: Network representation. The lighter node has more connections, considered the most prominent node.

In the network (see Fig. 3.4) the lighter node has the most connections in the network, making that the most active node in the network. It is a 'connector' or 'hub' in this network. The degree of a node is given by

$$k_i = \sum_{i \in N} a_{ij}, \tag{3.1}$$

and tells how many connections the node has to other nodes.

The degree of a node equals the number of edges connected to it. The statistical characterization of real networks displays a large number of node degrees, k, and the appearance of hubs, nodes with large degree. An undirected network with N nodes and L connections is characterized by an average degree $\langle k \rangle = \frac{2L}{N}$ (where $\langle \ldots \rangle$ denotes the average).

Despite the wide range of application, complex networks have developed to the characterization different topological networks, undirected and directed, unweighted and weighted. The techniques firstly applied to undirected and unweighted networks are lately adapted to weighted and/or directed networks. Topological properties have a very strong influence on propagation of knowledge and disease, as well as on robustness and vulnerability [Albert and Barabási, 1999; Dorogovtsev and Mendes, 2003]. Despite the importance of topological issues, weighted analyzes characterize the heterogeneity of weights and non-trivial correlation [Barrat et al., 2004a; de Montis et al., 2007].

Directed edges are considered when the edge from node i to node j ($i \rightarrow j$) is different of the edge from node j to node i ($j \rightarrow i$), see Fig. 3.3 middle. Many real networks are also weighted networks, in the case of social networks it is often relevant to assign a weight (strength) to each edge (see Fig. 3.3 right), measuring how good or strong is a relationship [Granovetter, 1973; Newman, 01 b; Marsden and Campbell, 1984].

In directed networks nodes have two degrees. The incoming degree k_{in} gives the number of connections to a node, and an outgoing degree k_{out} denotes the number of connections that start from the node and points to other nodes. The out degree of a node is the number of outgoing links, is given by:

$$k_i^{out} = \sum_{i \in N} a_{ij}. \tag{3.2}$$

The in degree of a node is the number of ingoing links, is given by:

$$k_i^{in} = \sum_{i \in N} a_{ji}. \tag{3.3}$$

On a directed network the total degree is given by the sum of the parts: $k_i^{total} = k_i^{in} + k_i^{out}$. The most basic topological characterization of a network is given by the degree distribution, $P(k)$. For each k the function $P(k)$ gives the average number of links that nodes with k connections have. For a directed network we can have $P(k^{in})$, $P(k^{out})$ and $P(k^{total})$. The degree distribution, $P(k)$, is a function describing the total number of nodes in a graph with a given degree:

$$p(k) = \sum_{j|k_j = k} 1. \tag{3.4}$$

This same information is often presented as the cumulative degree distribution:

$$P(k) = \sum_{k' \le k} p(k') \tag{3.5}$$

3.2.2 Length, Diameter and Shortest Path

Shortest paths play an important role in the transport and communication within a network. Suppose one needs to send a data packet from one computer to another through the Internet: the geodesic provides an optimal path way, since one would achieve a fast transfer and save system resources [Newman, 2003b]. For such a reason, shortest paths have also played an important role in the characterization of the internal structure of a graph [Wasserman et al., 1994; Scott, 2000].

The path between two nodes is the sequence of edges that one needs to take to go from one node to another. The shortest path to go from node i to j is named geodesic. The distance d_{ij} is the length of the geodesic from node i to node j. The maximum value of d_{ij} for every i and j is called the diameter of the graph, and will be indicated in the following as $diameter$ or *d*.

A measure of the typical separation between two nodes in the graph is given by the average shortest path length, also known as characteristic path length, defined as the mean of geodesic lengths over all couples of nodes [Watts, 2003b; Wasserman et al., 1994]:

$$L = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} 1/d_{ij}. \tag{3.6}$$

The average length and diameter is significantly different from network to network, and to the existing network models (see section 3.3).

3.2.3 Betweeness Centrality

So far we studied the local interactions of a node and its neighbors. But which are the prominent nodes on the global pictures? And how important are they when concerning information flow?

Common wisdom in personal networks is ”the more connections, the better,” although this is not always true. Some connections can be more crucial than others – bridging to otherwise disconnected nodes. On Fig. 3.5 the grey node has less connections than the node in the middle of the hexagon, but it is the ”connector”, bridging to the nodes on the right, which would be otherwise disconnected. She connects only those who are already connected to each other. The grey node plays a ’broker’ role in the network. It is central of information flow, like news and gossip in a social network, and plays a powerful role in the network. A node with high betweenness has great influence and power over the flows.

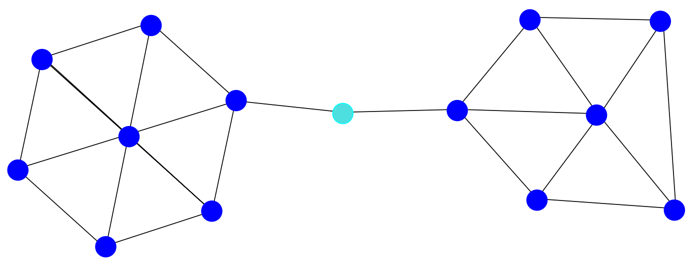


Fig. 3.5: Network representing betweeness centrality. The middle litter node has a higher betweeness, since all the information or traffic between the group in the left and the one in the right must go through that node.

Betweenness is one of the standard measures of node centrality, originally introduced to quantify the importance of an individual in a social network [Scott, 2000; Marsden, 1982]. Given by:

$$g(i) = \sum_{j \neq v \neq k} \frac{\theta_{jk}(i)}{\theta_{jk}}, \tag{3.7}$$

for θ_{jk} number of shortest path from j to k and $\theta_{jk}(i)$ is the number of shortest paths from j to k passing through the node i. The betweenness can be applied also to the edges, to measure the flow of information that crosses an edge. It is defined as the number of shortest paths between pairs of nodes that run through that edge:

$$g(\omega) = \sum_{j \neq v \neq k} \frac{\theta_{jk}(\omega)}{\theta_{jk}}, \tag{3.8}$$

for θ_{jk} number of shortest path from j to k and $\theta_{jk}(\omega)$ is the number of shortest paths from j to k passing through the edge ω.

The betweeness centrality differences from degree centrality for playing not the most central node, but the brokerage role or gatekeeper [Burt, 1995], also known in economics has the node with higher social capital [Portes, 1998]. To our knowledge this (quantitative) measurement has been applied only once on tourism literature and showed to be quite useful as an indicator on multidestination drive tourism [Shih, 2006]. Like in technologic networks, also on a network of tourist destinations the betweeness stands for the critical information that can be concentrated on that high betweeness node or edge. This critical destinations is the one where there is a very high probability that tourists cross or stop, being in this way an important intermediary.

One of the standard algorithms to calculate the shortest path is the Dijkstra's algorithm described below [Dijkstra, 1959]. Recently a fast algorithm for unweighted network was proposed by Brandes [2008]. Follows the main steps of the Dijkstra's algorithm, notice that for a weighted graph the distance between two nodes nodes can also be referred to as weight:

1. Create a distance list, a previous node list, a visited list, and a current node.
2. All the values in the distance list are set to infinity except the starting node which is set to zero.
3. All values in visited list are set to false.
4. All values in the previous list are set to a special value signifying that they are undefined, such as null.
5. Current vertex is set as the starting node.
6. Mark the current node as visited.

7. Update distance and previous lists based on those node which can be immediately reached from the current node.

8. Update the current node to the unvisited vertex that can be reached by the shortest path from the starting node.

9. Repeat (from step 6) until all nodes are visited.

The Dijkstra algorithm was applied on section 8.5 to measure the amount of traffic information on a given destination.

3.2.4 Clustering Coefficient

Initially, sociologists belived that the basic units of social networks were dyads, that is, the relations between two people would be the basic relational structure of the society. In this sense, the relations between the individuals on a group would connect on a more or less random way. Later the focus of analysis for the social networks would be the triads, of format triangular, also known as transitivity [Wasserman et al., 1994]. On the triads relation it is observed that two people having a friend in common also tend to know each other. These two people have, in this way, more possibility of knowing each other and of being part of the same group.

The social group or acquaintances were later studied with cliques [Wasserman et al., 1994], representing circles of friends or acquaintances in which every member knows every other member. This inherent tendency to clustering is quantified by the clustering coefficient [Strogatz, 2001], on the first model of social networks having a clustering close to the observed real-world networks (see section 3.3.3). Let us focus first on a selected node i in the network, having k_i edges which connect it to k_i other nodes. If the first neighbors of the original node were part of a group, there would be $\frac{k_i(k_i-1)}{2}$ edges between them. The ratio between the number E_i of edges that actual exist between these k_i nodes and the total number $\frac{k_i(k_i-1)}{2}$ gives the value of the clustering coefficient of node i:

$$c(i) = \frac{E}{k_i(k_i - 1)}. \tag{3.9}$$

The clustering coefficient of the whole network is the average of all nodes C_i's. In a random graph, since the edges are distributed randomly, the clustering coefficient is $C = p$, see section 3.3.1. However, it was Watts and Strogatz who first pointed out that in most, if not all, real networks the clustering coefficient is typically much larger than it is in a random network of equal number of nodes and edges. It is also observed by Newman and Watts [1999] that social networks have a much higher clustering than nonsocial networks. Clustering coefficient for weighted networks is introduced on section 8.2.

3.2.5 Assortativity

Assortativity is for a long time studied in social network analysis, where individuals having many connections tend to be connected with other highly connected individuals. To measure the correlation on the network over degree, one may also study the average nearest-neighbors degree. This measures the tendency of node i to be connected to nodes with the same degree,

$$k_{nn}(i) = \frac{1}{k_i} \sum_{j \in N_i} a_{ij} k_j, \tag{3.10}$$

which averaged over degree is:

$$k_{nn}(k) = \sum_{k'} k^{'} P(k'|k). \tag{3.11}$$

For a directed and weighted network the degree-degree correlations are introduced on section 8.1.

Degree-degree correlations are related with the concepts of assortativity and dissassortativity. Assortativity refers to a preference that nodes have to attach to other nodes that are similar or different in some way. This measurement is closely related with a typical behavior of many real-world networks. For example, in social networks, nodes with more connections tend to connect with other highly connected nodes. This tendency is referred to as assortativity. On the other hand, technological and biological networks typically show dissortativity, as high degree nodes tend to attach to low degree nodes. One of the measurements of assortativity/dissassortativity

Tab. 3.1: Assortativity of some real-world networks.

network	k_{nn}
e-arvix	assortative
company directors	assortative
internet	dissassortative
www	dissassortative
Protein interaction	dissassortative
neural network	dissassortative

are the degree-degree correlations on equation 3.11, for which the degree-degree correlations increase/decrease, see Fig. 3.6.

The assortative patterns of a variety of real world networks have been examined. On Tab. 3.1 the technological and biological networks all appear to be disassortative, while social networks are assortative.

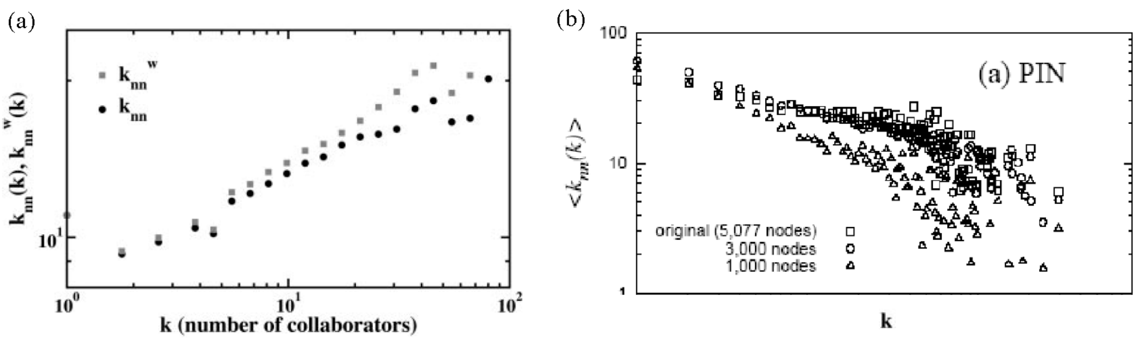


Fig. 3.6: Degree-degree correlations on a social network (e-print archive coauthorship network [Newman, 2001]) and on a biologic network PIN (protein interaction network [Lee et al., 2005]).

On this thesis the degree-degree correlations are analyzed for two networks on section 8.1 and 9.3, besides important properties of the world tourism networks, also a structural difference between the airports networks and the tourism network were found (see section 9.2). Based on those results is reconsidered a possible generalization of dissassortativity of travel and tourism sector, proposed on section 9.3.

3.2.6 K-Core

Another approach to define cohesive subgroups is the k-core, introduced by Seidman [1983], a subgraph in which each node is connected to at least a minimum number, k, of the other nodes in the subgraph, requiring a number of connections that must be present from each nodes to others within the subgraph [Dorogovtsev et al., 2006]. In this way are selected the most interconnected nodes globally, in contrast with the local interconnections depicted by the clustering algorithm.

K-core decomposition is a strong visualization tool for large networks [Alvarez-Hamelin et al., 2006b,a], and by studying the characteristics of the k-core nodes, it shows the central nodes most important on the evolutionary process of the network [Wuchty and Almaas, 2005]. The procedure is, by given a network, nodes with less than k connections are removed from the graph, recursively. These results in a series of sub networks that gradually reveal the globally central region of the original network.

It is interesting to note that the notion of k-cores has been recently used in biologically [Wuchty and Almaas, 2005; Wachi et al., 2005] related contexts, where it was applied to the analysis of protein interaction networks or in the prediction of protein functions. A further interesting application in the area of networking has been provided by Alvarez-Hamelin et al. [2006a] where the k-core decomposition is used for filtering out peripheral Autonomous Systems (AS) in the case of Internet maps (see section 3.4.1). Also in tourism Zach et al. [2008] approaches how possibly k-core can be a tool to find tourist destinations that are most active for the functioning of a region.

Hence, a k-core is a subgraph in which every node is a neighbor to at least k nodes, representing groups of a graph in which interesting nodes will be found [Seidman, 1983; Bollabas, 1984; Goltsev et al., 2006]. Follows two measures related with the k-core are graphically represented on Fig. 3.7.

On section 8.5 the k-core is analyzed for the world tourism network, along with the following two measurements for k-core characterization:

Definition 3.1 *A vertex i has coreness c if it belongs to the c-core but not to $(c+1)$-core. We denote by c_i the coreness of vertex i, and*

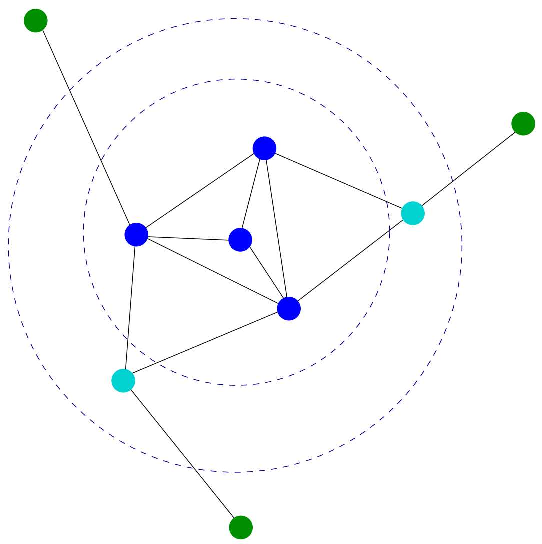


Fig. 3.7: K-core representation. Each node color represents a coreness and each circumference represents a k-core.

Definition 3.2 *A Shell C_c is composed by all the vertices whose coreness is c. The maximum value c such that C_c is not empty is denoted c_{max}. The k-core is the union of all the shells C_c with $c \geq k$.*

3.3 Network Models

The motivations for the development of a network theory were introduced on section 3.1.1 and the principal measures on section 3.2. On this section the main models of network evolution are presented as well as real-world applications, ranging from social to technologic networks, for which networks are relevant models.

3.3.1 Random Graph

The model of Erdös and Rényi [1960], that presents a random network, was the first model of social networks and its dynamics. It represents the simplest model of networks [Erdös and Rényi, 1959, 1960; Solomonoff and Rapoport, 1951; Gilbert, 1959]. The term random graph refers to the disordered nature of the arrangement of links between different nodes. Starting with N disconnected nodes, Erdös-Rényi (ER) random graphs are generated by connecting couples of randomly selected nodes, with no multiple connections, until the number of edges equals K [Erdös and Rényi, 1959]. For large number of nodes (N), and fixed $\langle k \rangle$, the degree distribution is well approximated by a Poisson distribution (Eq. 3.12). For this reason ER graphs are also called Poisson random graphs.

Gilbert [1959] proposed another model for random graphs, each pair of nodes connects with a probability p, such that $0 < p < 1$. The properties of the random graph vary with the probability p, with a critical change at $p_c = \frac{1}{N}$ corresponding to $\langle k \rangle_c = 1$. An interesting result on the structural properties of random graphs [West, 1995; Bollobás, 1985] is that:

- if $p < p_c$, the graph has no component of size greater than $O(\ln N)$, no component with more than one cycle;
- if $p = p_c$, the largest component has size $O(N^{\frac{2}{3}})$;
- if $p > p_c$, then graph has a component of O(N) with a number of O(N) of cycles, and there is no other component with more than $O(\ln N)$ nodes and more than one cycle.

For large N and fixed $\langle k \rangle$ the degree distribution is well approximated by a Poisson distribution:

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \tag{3.12}$$

Based on the definition of random graphs, they are uncorrelated graphs, since the edges are connected to nodes not considering their degree. Therefore the $P(k^{'}|k)$ and degree-degree correlations, k_{nn} are independent of k. Concerning the connectedness of the graph, for a $p \geq \frac{\ln N}{N}$ almost all the graphs with a given N and p are totally connected [Erdös and Rényi, 1959], and the diameter varies in a small range diameter $d \sim \frac{\ln N}{ln(pN)} = \frac{\ln N}{\ln \langle k \rangle}$ [Bollobás, 1985]. The average shortest path length between every two nodes has the same behavior has the diameter, $l \sim \frac{\ln N}{\ln \langle k \rangle}$ [Barrat and Weigt, 2000; Watts, 2003b; Bollobás, 1985; Watts and Strogatz, 1998]. The clustering coefficient of a random graph equals the probability of having a link between two nodes, $C = p = \frac{\langle k \rangle}{N}$. The ER have a very small clustering for graph with very large N.

3.3.2 Generalized Random Graphs

To better fit real data some generalizations of the random graph model were proposed. The configuration model is a generalization of the random graph (ER model), where a generic non-Poison degree distribution is given. A configuration model introduced by Bender and Canfield [1978] allows to generate a graph with $P(k)$, the degree sequence is given by a sequence of N integers $D = \{k_1, \cdots, k_N\}$ and $\sum_i k_i = 2K$, K is the total number of links. For large N the degree distribution tends to $P(k)$. The algorithm to generate the configuration model assigns to each node i a number of half-edges equal to the expected degree (from the D semble), the edges are added randomly, with uniform probability, to pair two half-edges together [Molloy and Reed, 1995, 1998].

A more general algorithm have been proposed by Bollabas [1984], the degree of nodes are independent identically distributed random integers from the desired distribution $P(k)$. This algorithm allows to obtain an ensemble of degree sequences for the given degree distribution, using a probability generating function [Bollabas, 1984]. The same authors have obtained an approximation for the shortest path length:

$$L = \frac{\ln(N/z_i)}{\ln(z_2/z_1)} + 1, \tag{3.13}$$

where $N \ll z_1$ and $z_2 \ll z_1$, for z_m an average number of neighbors at distance m.

Another proposed method for generating the generalized random graphs, but with a given expected degree sequence, was proposed by Chung and Lu [2002] and where its proofed that $L = \ln N / \ln \widetilde{d}$, for $\widetilde{d}$ equal of the sum of the squares of the degrees. The clustering coefficient in the configuration model is given by Bornholdt and Schuster [2003]:

$$C_{RG} = \frac{\langle k^2 \rangle}{N} \left[\frac{(\langle k^2 \rangle - \langle k \rangle)}{\langle k \rangle^2} \right]^2, \tag{3.14}$$

equals the ER clustering times another factor. Although for large $N \to \infty$ the clustering C vanishes.

3.3.3 Small-World Network Model

The previous presented models are both useful idealizations, but many real networks lie somewhere between the extremes of order and randomness. The small-world model was proposed by Watts and Strogatz whom conjectured that the same two properties short paths and high clustering would together fit for many natural and technological networks. Yet the network is much more highly clustered than a random graph, in the sense that if A is linked to B and B is linked to C, there is a greatly increased probability that A will also be linked to C (a property that sociologists call transitivity).

Watts and Strogatz studied regular lattice, the model starts with a ring lattice of N nodes, see Fig. 3.8, each symmetrically connected to its $2m$ nearest neighbors, where the total number of edges is $K = mN$. Then, for every node, each edge connected to a clockwise neighbor is then rewire by a shortcut edges added between random pairs of nodes, with probability p per link on the underlying lattice. Therefore the edge is preserved with a probability $(1 - p)$.

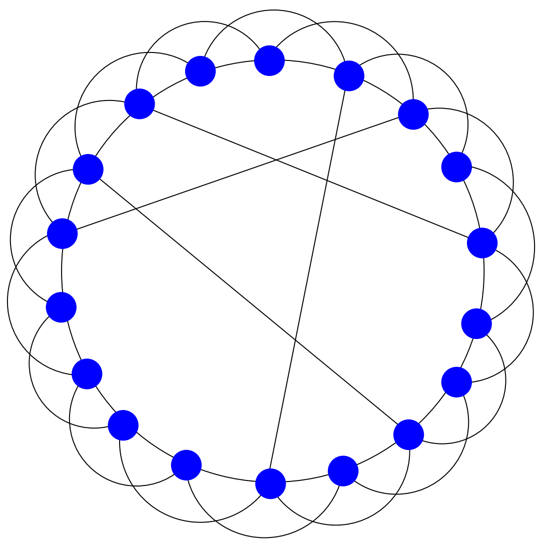


Fig. 3.8: Model of a small-world network.

Methods to obtain small-world networks based on adding new edges instead of rewiring have been proposed [Davidsen et al., 2002; Newman and Watts, 1999; Monasson, 1999]. The small-world model stimulated a grows on networks, where the structural properties of shortest path length and clustering have been studied, numerically and analytically for a given N and p [Barrat and Weigt, 2000; Newman and Watts, 1999; Barthélemy and Amaral, 1999].

The shortest path length, $L(p)$, has a drop when p is slightly higher than zero. In this way the small-world effect appears due to the rewiring to more distant nodes, making short cut links to otherwise distant nodes. The small-world property emerges on a logarithmic behavior. By other hand, the clustering C behaves with a linear drop, when each edge is rewired from the clustered neighborhood to another. The edges are not rewired with probability $(1-p)$, so two neighbor nodes that were initially connected, still connected with probability $(1-p)^3$ [Barrat and Weigt, 2000], therefore:

$$C(p) \sim \widetilde{C}(p) = \frac{3(m-1)}{2(2m-1)(1-p)^3}. \tag{3.15}$$

The $\widetilde{C}(p)$ is the average number of connection between neighbor nodes over the possible number of connections between the neighbors of a node. Barrat and Weigt [2000] analytically obtained the degree distribution for the small-world model for intermediate values of p:

$$P(k) = \sum_{i=0}^{min(k-m,m)} \binom{m}{i} (1-p)^i p^{m-i} \frac{(pm)^{k-m-i}}{(k-m-i)!} e^{-pm}, \tag{3.16}$$

for $k \geq m$, and equals zero for k smaller than m.

3.3.4 Scale-Free Networks

ER are the category of graphs most studied, although they do not produce most of the properties observed in real-world networks. The degree distribution in real-world networks is observed to follow a power-law, in contrast to the Poisson degree distribution obtained on the random graphs. Therefore the following models were stimulated to obtain a power-law, see sections 3.3.4 and 3.3.4.

Static Model

The first proposed models to obtain a power-law degree distribution were obtained by using the configuration model [Chung and Lu, 2002; Goh et al., 2001], see section 3.3.2. [Newman, 2003b] derived the clustering coefficient to a random graph with a power-law degree distribution, based on equation 8.8, $C \sim N^{(3\gamma-7)/(\gamma-1)}$, where the clustering tends to zero for $\gamma > 7/3$ and increases with the system size for $\gamma < 7/3$. On the former case the clustering emerges because there is more than one edge between two nodes with a common neighbor. Also for random graphs with a power-law degree distribution the shortest path length, $L \sim \log N$ if $\gamma > 3$. Conversely for $2 < \gamma < 3$ L is $O(\log \log N)$ and the diameter $O(\log N)$ [Chung and Lu, 2002; Cohen and Havlin, 2003].

A model based on the assumption that each nodes has a different weights (size) [Goh et al., 2001] $w_i = i^{-\alpha}$, where $i = 1, \cdots, N$ is the node sequence and α a given parameter in the range $[0, 1]$. A link between two nodes, i and j, depends on their normalized probability $w_i / \sum_l w_l$ and $w_j / \sum_l w_l$. The process is repeated until mN links are made in the system, so that $\langle k \rangle = 2m$. This model equal a random graph for $\alpha = 0$, otherwise its obtained a power-law degree distribution $P(k) \sim k^{-\gamma}$, for $\gamma = 1 + 1/\alpha$. For $\alpha \in [0, 1[$ implies $2 < \gamma \leq 3$ [Goh et al., 2001].

Caldarelli et al. [2002] proposed a model, *fitness model*, that starts with a given number of nodes N and assigns to every node, i, a fitness parameter, η_i, from a distribution $\rho(\eta)$. For every two nodes the probability of connection is given by a distribution $f(\eta_i, \eta_j)$, generating power-law distributions for many fitness distributions. For the uniform distribution $f(\eta_i, \eta_j) = 0 \; \forall i, j$.

Evolving Model

The models presented on the previous section assume a fixed number of nodes and degree distribution. Although, most networks in nature change over time, on the number of nodes, and develop on a dynamical process. The evolving models integrate a mechanism to simulate the evolution process of the network. The most popular self-organizing mechanism of networks is the one proposed by Albert and Barabási [1999]. In contrast with all the expectancies, it was found that most of the real networks display power-law shaped degree distribution:

$$P(k) \approx k^{-\theta} \tag{3.17}$$

By analyzing large networks researchers found that the structural principles of large real-world networks were based on a kind of preferential attachment. New members are added over time and the attachment prefers well connected members. This principle of ”preferential attachment” leads to relevant properties that have to be taken into account when analyzing and simulating systems with network concept. The recent term ”scale-free” has a new proposed constructor of real world networks comes as a better fit of real networks, using preferential attachment as a construction principle to model the networks.

The Barabási-Albert (BA) model was inspired on the World Wide Web growth [Albert et al., 1999]. It describes an evolving growth with preferential attachment and its evolving rules. The model integrates the idea that high degree nodes have a higher probability to be connected to the new nodes. Follows the algorithm to generate this mechanism:

1. Start with m_0 completely connected nodes;
2. Initialize a linear array where each node i of the network is presented k_i times. [At this step $k_i \equiv m_0 - 1, \forall_i$, and the array size is $m_0 \times (m_0 - 1)$];
3. At each step add a node to the network, and randomly choose m elements of the array of the previous step, to which the new node will connect. (To avoid multiple connections, if the same node is chosen more than once, then choose another random element until there is no repetition);
4. Update the array by adding to it m new entries corresponding to the new node, and another m entries corresponding to each selected node in the previous step;
5. Repeat step 3 until the desired network size N is reached.

The probability that a new node connects to an existing node is linearly proportional to the actual degree of the node i:

$$\prod_{j \to i} = \frac{k_i}{\sum_l k_l}. \tag{3.18}$$

The average distance for the BA model is smaller than in the random graph and increases logarithmic with the total number of nodes, N, and the clustering coefficient decays with the system size $C \approx N^{-0.75}$. Random graphs have a slower decay with $C \approx N^{-1}$. While for small-world models C is independent of N.

The BA model to explain the existence of power-law distributions in a wide range of applications. A power-law describes sizes distributions on many enormous complex physical and biological phenomenon, having an earliest result developed by [Simon, 1955; Bornholdt and Ebel, 2001; Newman, 2003b] about the sizes distribution of cities. Therefore, a great number of new models have been proposed to better fit empirical data.

Krapivsky et al. [2001] developed a directed network mechanism based on the BA model, where the in-degree and out-degree distributions of a growing network are modeled. The network is built based on the creation of new nodes which each immediately attach to a pre-existing node, and the creation of new links between pre-existing nodes. It generates degree-degree correlations. It overcomes a drawback of the BA model, the absence of degree-degree correlations.

Many generalizations of the BA model have been presented on the literature, based on nonlinear preferential attachment, dynamical edge rewiring, fitness models, etc. Dorogovtsev-Mendes-Samukhin (DMS) network model [Dorogovtsev et al., 2000] adds a linear preferential attachment of the form

$$\prod_{j \to i} = \frac{k_i + k_0}{\sum_l (k_l + k_0)}, \tag{3.19}$$

with $m_0 < k_0 < \infty$ [Dorogovtsev et al., 2000]. For $k_0 = 0$ it is the BA model. Otherwise m_0 represents the initial attractiveness of the node. For the attachment probability is obtained a power-law with exponent $\gamma = 3 + k_0/m$, so that $\gamma \in [2, \infty[$.

Krapivsky showed that the attachment probability is necessarily linear to obtain a power-law degree distribution [Krapivsky et al., 2001]. Another drawback of the BA model is that it does not allow to add or drop new edges once the network is created. Therefore some model were proposed that allow that nodes born and die or rewire [Dorogvotsev and Mendes, 2000; Krapivsky et al., 2001; Dorogovtsev et al., 2001; Tadic, 2001]. In real-world networks it can be that for example a more recent paper is more cited, or that for an older webpage is easier to have more links. Therefore new models try to better fit the real-world networks and their structural properties.

3.3.5 Network Model of Weighted Networks

There are some applications of the weighted models described in section 3.3.5, to market investment [Garlaschelli et al., 2005] and epidemic spreading [Yan et al., 2005].

To model a weighted network not only the process of the connection are needed but also to model the grows on the weights of the connections. The strength of each node also varies with the weights. A model can simply start with a random graph with a given degree distribution $P(k)$ with the weights distributed also given a distribution, $Q(w)$, which influences the final strength distribution $R(s)$ [Dorogovtsev and Mendes, 2004].

Yook-Jeong-Barabási-Tu (YJBT) proposed a model where the topology of the network grows with preferential attachment like the basic BA model, and the weights are assigned in relation with degree, $w_{ji} = \frac{k_i}{\sum_{i'} k_{i'}}$. Other model proposed different weight assignments, like $w_{ij} \propto (k_i k_j)^{\theta}$, $w_{ij} \propto max(k_i, k_j)$ and $w_{ij} \propto min(k_i, k_j)$. These models are based on growing topologies, and assign weights to the edges. A drawback is a missing dynamical grows of weights since dynamical reinforcement of edges is a common properties of real-world networks.

To overcome the dynamic of weights some models were presented. Barrat-Barthélemy-Vespignani (BBV) has a weight reinforcement mechanism related with the network growth [Barrat et al., 2004b]. The definition of the model is based on two coupled mechanisms: the topological growth and the weights' dynamics.

Starting from an initial seed of N_0 vertices connected by links with assigned weight w_0, a new vertex n is added at each time step. This new site is connected to m previously existing vertices, choosing preferentially sites with large strength; i.e. a node i is chosen according to the probability

$$\prod_{n \to i} = \frac{s_i}{\sum_j s_j} \tag{3.20}$$

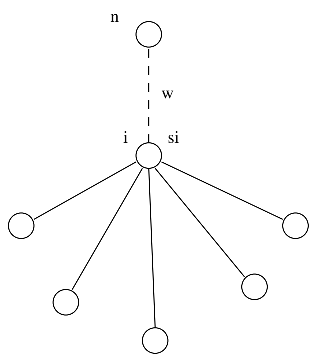


Fig. 3.9: Illustration of the construction rule. A new node n connects to a node i with probability proportional to $s_i / \sum_j s_j$. The weight of the new edge is w_0 and the total weight on the existing edges connected to i is modified by an amount equal to δ_i.

This rule, of strength driven attachment, generalizes the usual preferential attachment mechanism driven by the topology, to weighted networks. Here, new vertices connect more likely to vertices which are more central in terms of the strength of interactions.

This method corresponds to the fact that new sites try to connect to existing vertices with the largest strength. This is a plausible mechanism in many real world networks. For instance, in the Internet new routers connect to routers that have larger bandwidth and traffic handling capabilities. In the case of the airport's networks, new connections are generally established to airports with a large passenger traffic. In contrast to the connectivity preferential attachment of the "rich get richer" type, the mechanism here relies on the importance of the traffic and could be more adequately described as "busy get busier".

The weight of each new edge is initially set to a given value w_0. The creation of this edge will introduce variations of the traffic across the network. For the sake of simplicity we limit ourselves to the case where the introduction of a new edge on node i will trigger only local rearrangements of weights on the existing neighbors $j \in \nu(i)$, according to the rule:

$$w_{ij} \rightarrow w_{ij} + \triangle w_{ij} \tag{3.21}$$

where in general $\triangle w_{ij}$ depends on the local dynamics and can be a function of different parameters such as the weight w_{ij} , the connectivity or the strength of i, etc. In the following we focus on the case where the addition of a new edge with weight w0 induces a total increase δ_i of the total outgoing traffic and where this perturbation is proportionally distributed among the edges according to their weights, see Fig. 3.9:

$$\triangle w_{ij} = \delta \frac{w_{ij}}{s_i}. \tag{3.22}$$

The network generated by the BBV model display power-law behavior for the degree, weights and strength.

Another model with dynamical increase of weights is the Dorogovtsev and Mendes [2004] model. While in the first model the high strength nodes attract new edges and afterwards the weights of the edges of these nodes are specifically modified, in the DMS model high weight edges increase their weights and attract new connections. That is, in one approach, the attachment is to strong vertices and in the other approach, the attachment is to heavy edges. This is a principal difference which can be related to distinct real situations [Dorogovtsev and Mendes, 2004]. The rules of the DMS model are the following. The growth starts from an arbitrary configuration of nodes and edges, e.g., from a single edge of weight 1. At each time step: (i) an edge is chosen with a probability proportional to its weight and its weight is increased by a constant δ; (ii) a new vertex is attached to both the ends of this edge by edges of weight 1. The distribution of edge weight, node degree and node strength of the resulting network are power-laws.

3.4 Real World Networks: Society and Nature

On this section are presented some of the studied networks of real world systems. Complex weblike structures represent a wide diversity of systems of high technological and intellectual challenge. For example, the airports network is best described as a complex network where the physical places are connected by flights; the Internet is a complex network of routers and computers linked by various physical or wireless links; institutional rules and gossip spread on the social network whose nodes are human beings and edges represent various social relationships; the Wold-Wide Web is an enormous virtual network of webpages connected by hyperlinks. These systems represent just a few of the many examples that have recently prompted the scientific community to investigate the mechanisms that determine the topology of complex networks.

Also the weighted networks exhibit complex statistical features with highly varying distributions and power-law behaviors. Models have been discussed on section 3.3.5. Correlations between weights and topology provide a complementary perspective on the structural organization of such systems, like is discussed on the thesis data analysis section 6.

3.4.1 Technologic Networks, World-Wide-Web and the Internet

Technological networks are defining and designing the distribution of commodity or resource on the information society, such as the internet, electricity or information. Many real-world technologic networks have been studied, like the electric power grid by Strogatz [2001] and Amaral et al. [2000], the airline routes [Amaral et al., 2000] and railways [Latora and Marchiori, 2002]. On technologic networks the weight is provided by traffic and information flow.

The World-Wide-Web (WWW) is the largest network of information currently available, formed by hyperlinks between different Web pages and documents (web pages), with edges being the hyperlinks (URLs) that point from one document to another. The network reaches a size around 1 billion nodes at the end of 1999 [Lawrence and Giles, 1998].

The interest in the WWW as a network has boomed after it has been discovered that the degree distribution of the webpages follows a power-law over several orders of magnitude Albert et al. [1999]; Kumar et al. [1999]. Since the edges of the WWW are directed, the network is characterized by two degree distributions: the distribution of outgoing edges, $P_{out}(k)$, signifies the probability that a document has k outgoing hyperlinks and the distribution of incoming edges, $P_{in}(k)$, is the probability that k hyperlinks point to a certain document. Several studies have established that both $P_{out}(k) \approx k^{-\gamma_{out}}$ and $P_{in}(k) \approx k^{-\gamma_{in}}$ have power-law tails, with $\gamma_{out} = 2.45$ and $\gamma_{in} = 2.1$ [Albert and Barabási, 1999].

Faloutsos et al. [1999] reported the first study on how the Internet map looks like. Usually the Internet is described at the level of autonomous systems, or at the level of the routers. The Internet is formed, by a large number of domains of different administrative control, such as the autonomous systems (AS). Each AS is divided into many subnetworks and the routers are responsible for receiving and forwarding data packets, through both the subnetworks and among different AS.

The clustering ranged between 0.18 and 0.3, while the random network value is $C \approx 0.001$ with similar parameters. The average path length of the Internet at the domain level ranged between 3.70 and 3.77 [Vázquez et al., 2002]. The clustering coefficient has also a power-law decaying $C_k \approx k^{-\beta}$ and exponent $\beta = 0.75 \pm 0.03$ [Vázquez et al., 2002].

3.4.2 Social Networks

A social network represents actors (individuals or social groups) and relationships of a variety of kinds (friendship, kinship, status, sexual, business or political) [Wasserman et al., 1994; Scott, 2000]. The quantitative analysis of social interactions goes back to the early 1920s [Freeman, 1979]. The cross-interaction of researchers form diverse disciplines as sociology, applied anthropology, social psychology and statistics has raised over the years a solid scientific discipline with its own textbooks [Wasserman et al., 1994; Scott, 2000], and specialized journals as Social Networks published by Elsevier.

Anthropology and sociology have brought an important contribution into social network. The wide implication of network structure on diverse contexts has brought different strands in the nowadays social network analysis. There is no complete agreement on the separation of different strands. Gestalt theorists, during the 1930s, stimulated dramatically social network analysis [Moreno, 1978; Lewin, 1936; Heider, 1946] working on topics related to cognitive and social psychology. Moreno [1978], mostly interested in the potential of group settings for therapeutic practice, influenced by the German sociology of his time (Max Weber rationalization), developed the sociogram - a representation of formal properties of social configurations. This devise allowed the visualization of information flow which facilitates the understanding of the structure, individuals influence on each others, and identifying leaders. The graphical representation brought new insight to understand many properties of the network.

Topology and set theory were brought to the analysis of structural properties of social space - group and its perceived environment [Lewin, 1936], which contributed to one more approximation of social analysis and analytical techniques. Other of the most important concepts of social network analysis is balance theory. It conceptualizes the congruence of interpersonal attitudes toward psychological balance [Heider, 1946]. Heider's work on attitudes and perceptions concerned with how a person's various attitudes towards others are brought into a state of balance.

Lewin influenced the pioneer work of Cartwright and of the Harary to bring graph theory (network theory in mathematics lexicon) to group behavior. The cross of sciences was a breakthrough. Cartwright and Harary [1956] developed a framework for the crucial work of Lewin, Moreno and Heider. An important finding was the decomposition of large groups into subgroups with similar relations, this has derived others approaches: cliques, clusters and blocks (see methods in Wasserman et al. [1994]).

Radcliffe-Brown [1930] brought French sociology to British anthropology, constructing methodological foundations to frame ethnography (Social Organization of Australian Tribes). He was attributed with developing a sophisticated functionalist theoretical framework. His remarkable work strongly influenced nowadays anthropology and social network development. Algebraic models and set theory has brought new tools to study the concept of role in social structure. White et al. [1976] persisted explorations of block modeling (see also Wasserman et al. [1994]).

A precursor of another analysis was the remarkable work done by Granovetter [1973]. It has brought to the lexicon of sociology the famous sentence "the strength of weak ties". The argument assets that social coordination does not arise from densely interlocking strong ties, but from the presence of weak ties (less effective connected ties). This argument has already influenced organizational studies network analysis to consider the influence of weak ties to the cluster they belong. Social network analysis is a broad term that incorporates a variety of methods and applications, yielding a research tradition that is beyond the scope of this review to summarize fully.

Many of the fundamental concepts (such as the small-world property) and tools currently used in the analysis of complex networks have their origins in sociometry, like the clustering coefficient [Holland and Leinhardt, 1970], or of the different measures of node centrality proposed in sociometry to quantify the social importance of a given individual in a network [Wasserman et al., 1994; Scott, 2000], like betweeness (see section 3.2.3). Some current problems in network analysis, as the characterization of a node by its relations, have also been raised in sociometric studies. Concepts as the role or the equivalence of individuals were developed to locate actors placed similarly in a social network with respect to their set of relations [Wasserman et al., 1994]. Yet other problems such as searchability in networks, has been started by sociological experiments [Travers and Milgram, 1969], and measures such as the integration and the radiality have been proposed to quantify the degree an individual is connected and reachable within a given network [Valente and Foreman, 1998].

Collaboration Graphs

Within the framework of complex networks there have been some attempts to characterize the social interactions in animals (association, aggression, submission, grooming) [Lusseau and Newman, 2004]; the networked memberships of football players, musicians, and movie actors [Arenas et al., 2004]; or the interactions of fictional characters, such as the personages of Victor Hugo's Les Misérables, Tolstoy's Anna Karenina or Shakespeare's plays [Stiller et al., 2003].

Among the most studied social networks studied are the networks of collaboration similar to that of the movie actors can be constructed for scientists, where the nodes are the scientists and two nodes are connected if the two scientists have written an article together. To uncover the topology of this complex graph, Newman [01 b] studied four databases spanning physics, biomedical research, high-energy physics and computer science over a 5 year window (1995-1999). The degree distribution of the collaboration network of high-energy physicists is a power-law with an exponent of 1.2 [Newman et al., 2002], while the other databases display power-laws with a larger exponent in the tail.

All those networks show small average path length but high clustering coefficient. For example, for the science network on astro-ph has 22029 papers and 16706 authors with a clustering coefficient of $C = 0.4146$. Like referred in section 3.2.4 the clustering in social networks is very high when compared with nonsocial networks. There is a very strong clustering effect in the scientific community, on the database astro-ph two scientists typically have a 42 percent or greater probability of collaborating if they have both collaborated with another third scientist.

It was first pointed out that in most, if not all, real networks the clustering coefficient is typically much larger than it is in a random network of equal number of nodes and edges. The social networks have a high clustering, while random graphs that have edges distributed randomly have clustering coefficient is $C = p$ (see section 3.3.1 and 3.3.3). The behavior of $C(k)$ shows that authors with a low number of collaborators are more probable to work within groups where all scientists collaborate together than authors with a large degree. For $k \geq 10$ the weighted clustering coefficient $C^w(k)$ (defined on section 8.2) is larger than $C(k)$. Hence authors with large number of collaborators tend to publish more papers with interconnected groups of co-authors, and so influential scientists form stable research groups to produce the largest part of their papers [Barrat et al., 2004a].

Additionally the authors collaboration network formed a "small-worlds" where randomly chosen pairs of scientists are separated by only a short path of intermediate acquaintances. Moreover it was obtained that the minimum distance, in terms of numbers of intermediate acquaintances, on the whole database studied. The distance between the pairs of scientists is about six, which means that are six degrees of separation in science, like in the larger world of human acquaintance.

In the case of scientific collaborations, Newman [Newman, 2003a] has proposed to define the weight of the interaction between every two collaborators i and j as $w_{ij} = \sum_p \delta_i^p \delta_j^p / (n_p - 1)$, where p goes over all the papers, δ_i^p is 1 if author i has contributed to paper p and 0 otherwise, and n_p is the total number of authors of the paper p. The function has a normalization factor $n_p - 1$ that takes into account the authors of a large collaboration know one another less well, on average, than authors from a smaller collaboration, and the strength s_i of node i is equal to the number of papers author i has coauthored with others [Newman, 2003a].

The collaboration network is also assortativity, for $k_{nn}(k)$ and $k_{nn}^W(k)$ growing as a power-

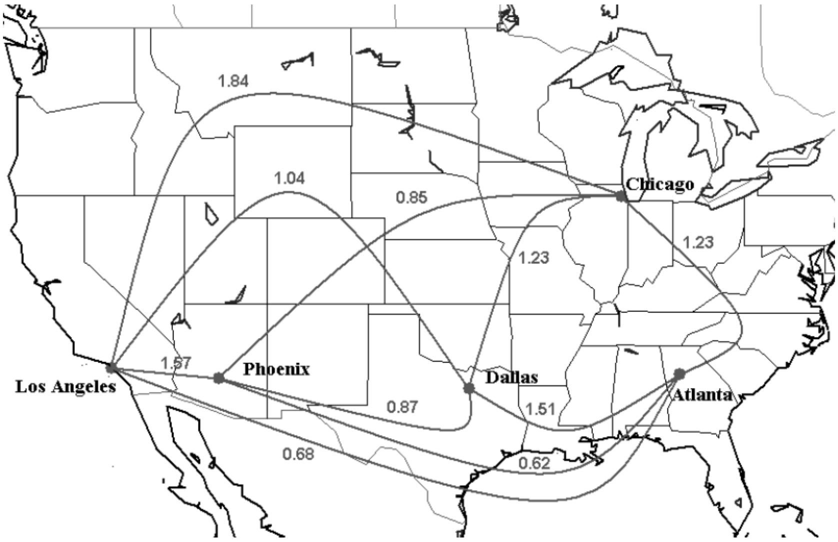


Fig. 3.10: Degree-degree correlations and clustering. Source: [Barrat et al., 2004a].

law as a function of k, therefore it has a behavior typical from social networks [Newman, 2003a] (see section 3.2.5).

The Web of Human Sexual Contacts

Many sexually transmitted diseases, including AIDS, spread on a network of sexual relationships. Liljeros et al. [2001]; Liljeros [2004] have studied the web constructed from the sexual relations of 2810 individuals, based on an extensive survey conducted in Sweden in 1996. The analyzed the distribution of partners was made over a single year, due to the sort period that edges exist, obtaining a power-law degree distribution for females with an exponent of $\gamma = 3.5 \pm 0.2$ and for males of $\gamma = 3.3 \pm 0.2$.

3.4.3 Air Transportation Network

The world air transportation network has been studied in the last decade, representing an example of a large infrastructure systems.

Barrat et al. [2004a] analyzed the International Air Transportation Association (IATA) [IATA, 2006] database containing the world list of airports pairs connected by direct flights and the number of available seats on any given connection for the year 2002. The resulting air-transportation graph comprises $N = 3880$ nodes denoting airports and $E = 18810$ edges accounting for the presence of a direct flight connection. The average degree of the network is $\langle k \rangle = 9.70$, and maximal degree 318.

The topology of the graph exhibits both small-world and scale-free properties as already observed in different dataset analyses [Li and Cai, 2003; Guimera et al., 2005]. In particular, the average shortest path length, measured as the average number of edges separating any two nodes in the network, shows the value $\langle l \rangle = 4.37$, very small compared to the network size N.

The degree distribution, on the other hand, takes the form $P(k) = k^{\theta} f(\frac{k}{k_x})$, where $\theta \approx 2.0$ and $f(\frac{k}{k_x})$ is an exponential cut-off function, that finds its origin in physical constraints on the maximum number of connections that a single airport can handle [Guimera et al., 2005; Dorogovtsev and Mendes, 2003]. The airport connection graph is therefore a clear example of heterogeneous network showing scale-free properties on a definite range of degree values.

The properties of a graph can be expressed via its adjacency matrix a_{ij} , whose elements take the value 1 if an edge connects the vertex i to the vertex j and 0 otherwise. The data contained in the previous datasets permit to go beyond this topological representation by defining a weighted graph that assigns a weight or value characterizing each connecting link. In the case of the airports network the weight w_{ij} of an edge linking airports i and j represents the number of available seats in flights between these two airports. The average numbers of seats in both directions was identical $w_{ij} = w_{ji}$ for an overwhelming majority of edges, therefore the network is considered undirected.

Barrat et al. [2004a] have studied the world-wide airport network as a weighted graph with the weights w_{ij} being given by the number of available seats on direct flight connections between the airports i and j . The degree distribution takes the form $P(k) = k^{-\gamma} f(k/k_x)$, where $\gamma \sim 2.0$ and $f(k/k_x)$ is an exponential cut-off function that finds its origin in physical constraints on the maximum number of connections that a single airport can handle [Guimera et al., 2005]. The probability that a vertex has strength s is heavy tailed and the weights are non-trivially correlated to the degree: the average link weight scales with the degrees of the two end-nodes as $\langle w_{ij} \rangle \sim (k_i k_j)^\theta$, with an exponent $\theta = 0.5$ [Barrat et al., 2004a].

The strength of vertices of degree k follows a power-law behavior $s(k) \simeq Ak^\beta$ with an exponent $\beta = 1.5 \neq 1$ [Barrat et al., 2004a]. This means that the larger is an airport, the more traffic it can handle. The weighted clustering coefficient $C^w(k)$ has much more limited variation in the whole spectrum of k. This implies that high degree airports have a progressive tendency to form interconnected groups with high traffic links, thus balancing the reduced topological clustering.

Since high traffic is associated to hubs, the airports network have a network in which high degree nodes tend to form cliques with nodes with equal or higher degree, the so-called rich-club phenomenon [Zhou and Mondragón, 2004]. The topological $k_{nn}(k)$ does show an assortative behavior only at small degrees. For $k > 10$, $k_{nn}(k)$ approaches a constant value, a fact revealing a uncorrelated structure in which vertices with very different degrees have a very similar neighborhood. The analysis of the weighted $k^w_{nn}(k)$, however, exhibits a pronounced assortative behavior in the whole k spectrum, providing a different picture in which high degree airports have a larger affinity for other large airports where the major part of the traffic is directed.

4 Challenge to Tourism Systems: Social Network Analysis

Network are all around us and network industries play a crucial role in modern life. Networks of friends, interorganizational nets, the internet, are just a few examples of real-world networks. The notion of networks as a dominant organizing principle to explain how the world works have been influencing over the decades scientists in many areas. The modern economy would be very much diminished without the transportation, communications, information, and railroad networks.

On this section we focus on the concept of networks on the travel and tourism industry. The first studied networks in tourism were interorganizational nets, more recently other forms of networks have been studied. The simple knowledge representation of the interdependency of organizations, or the way destinations are interconnected through transportation or product similarity, starts with the understanding of the system as a whole. This path of knowledge representation starts with the categorization of entities, organizations or people, and the structure of their relationships.

The study of network provides rich insights into the way interconnected systems are constructed influence the dynamics and communication processes in many kinds of socio-economic systems. Tourism destinations are one of the most networked socio-economic system. It is characterized by multiple actors, sharing a common product and mutually dependent on a network system. Due to the interdependency among actors, their connections are considered a main competitive factor for a tourism destination [Scott et al., 2008b; Giménez-García et al., 2007; Morrison et al., 2004].

Tourism is a highly networked sector where clusters of organizations within a destination, as well as networks of cooperative and competitive organizations linking destinations [Lemmetyinen and Go, 2008; Crotts and Aziz, 1998; Yuksel and Yuksel, 2004; Saxena, 2005]. The paradigmatic abstraction of the structural relationships assumes that individuals act depending on the pattern of relations. This is in contrast with the classical organizational and economic concept based on individual properties. For business, networks are successful vehicles of cooperation and leverage, organizations connect themselves through partnerships and clusters forming communities.

In the past decade the concept of networks in tourism has been applied to an increasingly number of areas to explain outcomes such as competition, volunteer tourism [McGehee and Santos, 2005], inter-firm relations, among others [Morrison et al., 2004; Pavlovich, 2003; Schelling, 1978; Harshaw and Tindall, 2005b]. The two stronger perspectives on the study of tourism network systems are interorganizational and multi-destination.

Without getting into a full record of the historical development we intent to understand the importance of the characteristics of this new structure of interdependences within the travel and tourism sector. In accordance with business and economic science [Burt, 1995], tourism also bases on the assumption that destinations are expected to depend on the success of partnerships or other forms of networking in a destination [Costa, 1996; Morrison et al., 2004; Pavlovich, 2003]. Relationships are seen as an important component of competitiveness. The interorganizational network on tourism cover several types of relationships, as discussed in section 4.1. While there is a growing interest in interorganizational networks, little has been done on multi-destination concept and the mapping of product dependency. The structure of product dependency and destination net is covered on section 4.2. By covering the literature on networks a diverse range of concepts and definitions is identified.

The term of networks, networking and clusters are confusing within the tourism literature, where the terms are found to be disparate inconsistent in terms of terminology and the concepts adopted, as well as used as everyday speech terminology. Shaw and Conway [2000] note that the term ”network” has been loosely applied in entrepreneurial research, and, as a consequence, suggests there has been a general failure to recognize that networks and networking are different constructs.

The semantics of definition are taken still further when Rosenfeld [2001] seek to distinguish networks from clusters; but, in so doing, they rely on restricted definitions of network, which perhaps simply emphasizes the difficulties of definitional precision.

A more consistent interpretation, however, can be gleaned from social network theory, for investigation into the networks in which small firms are socially embedded [Baht and Milne, 2007]. Curran and Storey [1993] (p. 13) argue that 'networks' are best seen "... primary cultural phenomenon...", that is "... as sets of meanings, norms and expectations, usually linked with behavioral correlates of various kinds": but they suggest it is "... the meanings, norms and expectations which are important..." and that the behavior correlates serve largely to indicate the kinds of social relations that are worth investigating.

Such varied conceptual starting points inevitably shape conceptual starting points inevitable shape the different ways of knowing 'networks' and 'networking' and impact on how they are defined. Nevertheless, there seems to be a level of agreement that it is not the existence of a network that in itself has the potential to generate benefits but rather it is the use of that network through the process of networking that actually brings about the gains for the networks' membership. The critical element to this distinction rests on an understanding on how the ideas and patterns of action develop among groups and individuals [Halme, 2001], suggesting the significance of the social dimensions to entrepreneur decision-making. In short, a network might simplistically be viewed as the structure and scaffolding that supports and contains networking, but this is rich in social meaning, texture and the relationship involved in the process of networking.

There is a mixed approach to what might constitute an appropriate definition of a network: for example, they have been described as an organized system [Szarka, 1990], or as a specific type of casual alliance [Knoke and Kuklinski, 1983], an informal association Thompson et al. [1991]; Harshaw and Tindall [2005a]. Informal networks can also represent the existing social relations between tourists [Murphy, 2001; McKercher and Fu, 2006]. Thus, while there appears to be agreement that networks consist of a series of directed and indirected ties from one actor to a collection of others [Smith-Ring, 1999], the level of agreement about the formality required to bind these ties is at variance.

There is a consensus that the network ties represent co-operative conduits linking firms to a set of persons, objects or events [Knoke and Kuklinski, 1983], or to other business firms, governmental bodies or organizations, persons or others [Smith-Ring, 1999], and that it is essentially these ties which lubricate the social relation between the network's actors. This also helps in the coordination of the network participants' political and economic lives [Thompson et al., 1991] for the exchange of ideas and information to their mutual benefit [Knoke and Kuklinski, 1983].

Lynch [2000] observes that, "... the term (networks) describes the interactions of the firm with the external environment, and offers potential insights into such areas as business relations, industrial organization, regional agglomeration, strategic management of small firms and the culturally induced outlooks and behavior of small firms."

The process of networking, then, can be defined as the activation of the actors, relationships, ties, inter-connections, conduits and content that has been framed within a network structure. It is possible to identify dimensions of networking to include: level of networking, networking proactivity and the strength of network ties. The reality, however, remains that it is not the existence of a network per se from which the benefits will accrue, but rather the use of that network through the process of networking that generates the desired outcomes [Shaw and Conway, 2000].

On the last few years research has also addressed the structural characteristics associated with the multiple tourist destinations. Without this realization, it is difficult for government planners and tourism service providers to develop appropriate facilities at a particular destination based on the network position of the destination within various tourist routes (see section 4.2). Network analysis can serve as an appropriate method for measuring that. The following section introduces the techniques and indicators of network analysis adopted in this study to address this issue.

4.1 Interorganizational and Governance Networks

Interorganizational networks have become a popular form of organization cooperative relationship, and increasing attention is being given to them by academics, analysts, and operators alike in the fields of hospitality and tourism [Petrillo and Swarbrooke, 2005; Scott et al., 2008a]. Issues of partnership and collaboration have come to the vanguard of tourism research [Yuksel and Yuksel, 2004; Hall, 2000; Scott et al., 2008b]. Within the context of the tourism sector, while it is recognized that networks serve multiple purposes, one such function is their role in assisting the formation of alliances to facilitate the packing of a series of related products, marketing collaboration [Wang and Fesenmaier, 2007; Larson, 2002; Fyall et al., 2001] or services at specific destinations.

Interorganizational networks is a phenomenon which is also being given attention in the management literature on business development and economics, which have become increasingly aware of how strategic alliances and network formation have become an essential feature of corporate power in modern economies [Axelsson and Easton, 1992; Burt, 1995].

From an interorganizational side, Lynch [2000] observes that, ”the term (networks) describes the interactions of the firm with the external environment, and offers potential insights into such areas as business relations, industrial organization, regional agglomeration, strategic management of small firms and the culturally induced outlooks and behavior of small firms.” Additionally Crotts [1998] refers the main purpose of business networks the increase of competitiveness of each firm.

At at generic level, a range of different types of networks are easily identified, and this maybe classified in various ways. For example, Colloquium [1998] devised a network classification, which is provided in Tab. 4.1. This particular classification is useful as a means for conceptualizing networks. While it is comprehensive, it does neglect the concept of the virtual organization that exists only through cyber-technology, for example, via the internet or e-mail; and, as a further omission, Halme [2001] notes that networks may also vary according to the type of organizational configuration that pertains.

Morrison et al. [2004] recognizes that a deeper appreciation of the content of network

Tab. 4.1: Network classification.

Classification	Description
Network membership	Diversity of actors (professional, user, social)
Nature of linkage	Formal versus informal
Type of exchange or transaction	Information, goods, friendship or power
Network function or role	Problem solving or idea generation
Network morphology	Size, diversity, density, stability of links
Geographical distribution of network	Balance between local, national and international members

relationships would increase understanding of the process of networking. It is described that social network are a set of morphological dimensions that consider both the pattern and the structure of the network, with its interactional dimensions, which are based on considerations of networks as a process. Thus, it has been argued that network analysis considers both the the structure and the process of the relationships that join individuals, groups and organizations [Granovetter, 1973], while it is the content of those relations that actually captures the meaning people attach to the relationship that are formed.

Lynch [2000] shows that network analysis has provided a mechanism by which one can explore aspects of nature and function of social and business networks, as well as the meanings, norms and expectations supporting those relationships in small hospitality enterprizes and those which may be considered as 'minority firms' [Ram, 1994]. Thus, the application of social analysis to the structuring of networks and the process of networking aims at identifying the pattern, content, meanings, motivations, expectations, norms, and nature of interactional relationships. This contributes to an improved understanding of socially embedded relationships, which is of particular significance as it has been demonstrated that market exchanges are embedded in, and defined by, more complex social process [Pavlovich, 2003].

The cooperation required for a network to operate successfully is perceived as essential for any tourist destination. For example, Lynch and Tinsleyand [2001] provide an empirical example of informal tourism-related networking in a rural destination on the west coast of Scotland. They investigated the networking process between the destination's small tourism business - including hotels, beds and breakfasts, gift shop, art and craft shop and grocery store - to demonstrate the role that co-operation played in the destination's development. Therefore, many studies emphasize the role of networking as an effective option in terms of mobilizing information and resources, and of developing a cooperative processes among tourism businesses [Gnoth, 2002; Buhalis, 2000].

Hence, in the 'open-ended' tourism setting, researchers emphasize that networking represents a crucial and effective option in terms of mobilizing information and resources, and of engaging in cooperative processes among tourism businesses [Buhalis, 2000; Gnoth, 2002; Scott et al., 2008b; Gretzel et al., 2008; Hall, 2000; Bramwell and Sharman, 1999; Araujo and Bramwell, 2002].

The conceptualization of 'the tourism system as a network of interacting service providers', Gnoth [2002] may provide an effective mechanism for bringing about community involvement, in particular through the selection of key stakeholders who represent and hub serving the various interests within society.

Palmer [1996] offers a public-private sector network example in findings from English district councils, driven by the rationale that attracting more tourists to regional tourist destinations can benefit not only the narrow financial objectives of tourism operators, but also the more diverse social objectives of the public sector.

Costa et al. [2007] examine the benefits of innovative networks and partnerships, carried out in Portugal, targeting sports and adventure tourism enterprizes, with an empirical research demonstrating that SMEs clearly assume an important role in economic development, coordinating much of their business through informal networks.

Networking has been recognized to leverage business in many specialized forms of tourism, like rural tourism, museums networks [Tufts and Milne, 1999], volunteer tourism, etc. Participation in volunteer tourism has a positive effect on social movement activities [Tufts and Milne, 1999; McGehee and Santos, 2005], the notion of network is a cooperative strength. On cooperation among museums has always been a part of everyday operations (such as exchange of artifacts and traveling exhibits). There is a significant amount of cooperation and networking among museums which at first appears to mimic interfirm network relations such as joint marketing and externalized operations. Moreover the networking concept also is refereed as a primary tool for event leverage and relationship development addresses in large international events, like the olympic games [O'Brien, 2006].

National and provincial organizations are the more established and formal networks servicing the sector's collective lobbying, literary, and research and development needs. Cooperation between museums and various governmental tourism agencies is at times just as important as networking among the former. For many, marketing initiatives beyond print media are only financially feasible with the assistance of joint publicity programs or subsidized tourism packages.

On policy networks Pforr [2006] studied how public, private, and nonprofit actors shape policymaking processes and policy outputs in the specific geographical and social formation of the Northern Territory of Australia. In the context of tourism with its vast variety of actors and complex web of interactions, the focus can be set on the participants in the policymaking process, their relationships, and the structural context in which these take place.

Network emerge on local, regional and international level as tourism becomes a predominant sector. The globalization intensification increased the ways of international and regional interdependence and governance, like the creation of United Nations and European Union. Tourism sector had an international dimension from the beginning, and the need for cross border cooperation came forward very early. The first International Congress of Official Tourism Traffic Associations was held in The Hague, in 1925. International cooperation was also broadened and intensified on the same premises. This led to the constitution of the International Union of Official Travel Organisations (IUOTO) in 1947.

Tourism, like the economy in general, will have to deal with the trend of globalisation, the growing power of international economic and market forces and the consequent decline in the ability of individual states to control their economies and of private corporations in limited geographic spheres of operation to dominate domestic markets. The impact on tourism is that more power will be obtained by the relatively small number of global travel and tourism networks achieving their globalisation not only through vertical and horizontal integration but through diagonal integration [Castells, 1996], economies of both scale and scope, and their huge investment in electronic databases and marketing.

In terms of non-governmental networks, a number of European or multinational associations are being formed to promote a particular theme issue relating to cultural heritage in rural areas. Europe of Traditions, for instance, brings together individual home owners and associations from England, France, Ireland, the Netherlands and Portugal who are interested in preserving local vernacular buildings by sensitively transforming them into high quality accommodation that reflect the history, hospitality and heritage of the area.

There are also calls to recognise fully the role of indigenous peoples in respect of protected areas [Beltran, 2000], and to develop international co-operation in protected areas across national boundaries [Sandwith et al., 2001]. The PAN (Protected Areas Network) Parks Initiative began in 1997. The idea of this initiative was ”to introduce a marriage between nature conservation and tourism on a European scale” [Hogan, 2000]. The initiative aims to put the economic value generated through tourism into the protection of Europe’s nature.

Tourism strategy networks can operate at any level. Some of the most successful to date have involved different forms of accommodation, such as the ”Gites de France” network or the ”Europe de Traditions”, which group together similar types of accommodation around a common theme. Others concentrate more on a specific tourism segment, for instance, the ”Pan Parks” [Hogan, 2000], aims to link up different national parks across Europe through a uniform vision of high quality tourism and strong conservation principles. Alternatively, different cultural or natural routes and itineraries could be explored, such as those promoted by the Council of Europe.

Hence, many case studies recognize networks as a catalyzer on marketing [Wang and Fesenmaier, 2007], policy-making [Dredge, 2006], knowledge transfer and innovation dissemination [Novelli et al., 2005].

4.2 ICT and multi-destination Networks

New forms of organizations arise in different cultural contexts, adapting to the new information era. As a matter of course evolve a common form of organizations interactions. Like interorganizational networks, international networks brings a growing interaction among cultures, which is intensified by new emerging channels of communication. A social, economic and technological networked interdependent economy is shaping a not planned increasingly transformation. As the WTO recognized in the 2000 report [WTO, 2000], two key forces, globalization and technology, are transforming the tourism sector into a dynamic economic force that has never been possible before [Smeral, 1998]. Therefore, tourism sector came up during a transition period in our society structure. Castells [1996] has referred that we are witnessing a point of historical discontinuity.

The emergency of information and communication technologies (ICT) plays the capability of hospitality enterprises to simultaneously compete and cooperate, within the destination and event network context. The effects of the former appear to be profound in that they seem to spread fast and change not only the methods of production, but what is being produced, and how, as well.

In particular, Breukel and Go [2008] seek to identify insights into how network stakeholders, both local and global deal with the key issue of disruptive innovation [Novelli et al., 2005], its impact on business processes [Giménez-Gracía et al., 2007] and the subsequent need for participation in the supportive context of ICTs to reduce transaction cost and improve service quality within a network environment. Authors such Buhalis and Main [1998]; Lang [2000] refer to disintermediation as a consequent phenomenon of Internet's use in the tourism sector. Breukel and Go [2008] refers to the impact on internationalization of destinations and tourism stakeholders as ”the tourism system as a network of interacting service providers rather than as a channel of distribution” as a new strategic perspective for tourism branding.

In the same way that tourism destinations can be conceptualized as a networked system, their electronic image can also be percept and used as a networking strength. Network analysis focuses on the structure of relations between given actors, and applies techniques to produce relevant indicators and results for studying the properties of the network as a system. Previously applied to relationship in destinations, the concept of networking to compete and market can now be applied to an e-destination. The use of new technologies can change the perspective of a destination and new challenges emerge.

For example, the use of mobile system as tour guides increases the amount of information obtained by the tourists and result in a longer stay at each tourism attraction [Kramer et al., 2005]. This is achieved because more integrated information is delivered and easy to access. Web pages, mobile devises and GPS are new ways of connecting, whether it is connecting organizations [Baggio, 2007], tourism attractions [Kramer et al., 2005], tourists [Ahasa et al., 2008], tourist transportation [Lumsdon, 2006], or any other dependency of the product chain. In this way, new technologies increase integration of tourism chain into a single product, whether it is at the destination or on its online image, e-destination.

In a tourism destination, companies have become conscious that in order to compete globally, they have to interconnect tightly with the surrounding local environment in order to become more efficient in their operation. The growing adoption and evolution of new communication channels places the internet on a primary importance for tourism destinations. The webspace is a new channel for branding tourist destinations and in this internationalize [Hawkins, 2004].

A common strategy online may benefit all the actors. Faloutsos et al. [1999] reported on the first study on how the internet map looks like. Internet and world-wide-web are valuable objects because modern society increasingly depends on large communication networks. The need for information spreading pervades our lives, and its efficient handling and delivery, is becoming one of the most important practical problems. Pushed by this practical need, developing a realistic e-destination network and understanding the basic mechanism of the formation of such networks are crucial for online positioning, cost optimization and efficiency.

The online image of a destination competes through interconnected hyperlinks and by connecting the destination to the global online market. Baggio [2007] mapped an e-destination showing that stakeholders have low connectivity, with a significant amount of disconnected organizations. This can show that stakeholders and public organizations still did not perceive the importance of being interconnected online. Destination management companies can probably play a hub role to connect otherwise disconnected organizations. The networked structure of a destination is discussed on section 9.3. The world-wide-web and the internet were modeled into a complex network, showing its dynamic evolution and topological structure, see section 3.4.1.

Like organizations, tourism attractions and resources are also linked into a common product. In this sense, a multi-destination network [Shih, 2006; Mathews, 2000] conceptualize a set of different destinations or sights as a single product. Network analysis has been proposed in tourism research to quantify the structure of a multi-destination that functions as a single product.

This product can be a set of different cities visited by airplane [Hwang et al., 2006], a grouping of different tourism attractions on a single city [Kramer et al., 2005] or a set of cities around a hub main city visited by car on a single trip [Shih, 2006]. This is achieved because more integrated information is delivered and easy to access. Web pages, mobile devises and GPS are new ways of connecting, whether it is connecting organizations, tourism attractions or any other dependency of the product chain. In this way, new technologies increase the integration of tourism attractions into a single product.

Multi-destination also benefits common products, that can be easier to market together. The grouping of cities around themes is particularly important in the South, as seen in the strong participation in the Arts Cities of Europe network. But also in other parts of Europe, cities link themselves, for example "Magic Cities Germany" (with ten participants, including Berlin, Dresden, Frankfurt, Hamburg, Cologne and Munich), or the Golden Triangle (Budapest, Vienna and Prague).

On the next section is discussed how network analysis improves our understanding of tourism interorganizational and multi-destination networks.

4.3 Network Theory and Tourism

On this section is shown the interplay between empirical evidences of the benefits of networks with the networks, both theory and methodology. On the theoretical part is reviewed how researchers relate the main theoretical results on networks with tourism systems. By other hand network analysis methods on tourism research are reviewed.

4.3.1 Social Network Analysis

During the last decade there has been more emphasis on the application of quantitative methods on the study of tourism networks [Shih, 2006; Pavlovich, 2003; Scott et al., 2008b, 2007; Miguéns and Mendes, 2008b; Baggio, 2007]. With the first work on network analysis quantifiers by Shih [2006], mainly on centrality (see section 3.2.1), betweeness (see section 3.2.3) and on quantifiers related with the theory of structural holes [Burt, 1995]. Shih [2006] showed that network analysis is an appropriate tool to understand the relational structure of a network of destinations on a given region. Lately on the interorganizational perspective cohesion, clustering, density and modularity were introduce by Scott et al. [2008b], that showed how network analysis brings a diverse set of methods. It can be important for comparing different destinations, on a way that future simulation of destinations can be performed to detect failures.

Pavlovich [2003] shows how the evolution and transformation of a tourism destination is related with the network theory, respectively the strength of weak ties [Granovetter, 1973] and the structural holes [Burt, 1995]. This researcher offered a strong conceptual contribution by showing how network theory has a theoretical contribution on explaining the evolution of linkages between organizations.

The collaborative nature of tourism competitiveness has been recognized in many case studies [Pavlovich, 2003; Scott et al., 2008b; Ritchie, 1999; Bhat and Milne, 2008]. Pavlovich [2003] shows how strong and weak ties [Granovetter, 1973] play different roles on a destination, the strong ties exchange mechanisms and weak ties informal opportunities [Werde et al., 2005]. Following the same idea the researcher showed that much of the marketing occurs through weak ties and that different nodes or organizations have independent weak ties that introduce new knowledge into the network. Each of those linkages has a potential value to the destination is compared to a *structural hole* [Burt, 1995] that bring new information into the networked destination.

In general Granovetter [1973] distinguished two types of ties on networks, the strong ties and the weak ties. The first ones identified on high clustered groups where relations diffuse acceptability and inclusion. By other hand weak ties are the connections outside high clusters which connect to other groups, knowing to be necessary to gain new ideas and opportunities from an external environment. On management literature the theory of structural holes [Burt, 1995] also states the importance of the connectors, or the nodes that connect otherwise disconnected groups, claiming its importance to entrepreneurship and diffusion of information.

Network theory is characterized by a constant interplay between sciences that are converting into a common framework or common theory. Many of the results obtained in sociology, like the strength of weak ties, or the six degree of separations were lately generalized by physicists (see section 3), into the science of complex networks. On the methodology section 5.1 is reviewed how network theory is interpreted has a methodology, theory or paradigm. The tourism systems benefit from network analysis in the same way, by also integrating network as a theory and a methodology. Therefore, the theoretical results from the strength of weak ties, six degrees of separation, structural holes or scale-free networks are observed phenomenons on many real-world networks. In this way, tourism systems benefit not only on using the methodologies of network analysis but also its theoretical results [Pavlovich, 2003; Baggio, 2007; Miguéns and Mendes, 2008b; Scott et al., 2008a].

Visualization programs and techniques [Batagelj and Mrvar, 2002; Borgatti et al., 2002] have also proved to be efficient methods to depict patterns within a network. The immediate understanding of small networks through a mapping can help researchers and managers. A destination that has an organization coordinating the communication among stakeholders can benefit of network mapping to more easier explain the importance of networking to the stakeholders. A drawback of visualization is that for networks with a considerable amount of nodes the mapping becomes too unclear. The amount of data needed to analyze relationships on a destination can be costly forbidden, and in this perspective information and communication technologies can provide new way of analyzing relationships. ICT's brought new potentialities to collect information, and for organizations to compete on the market place (see section 4.2).

4.3.2 Centrality and Density

One of the central measurements of network theory is the concept of centrality, discussed on section 3.2.1. In accordance with the general science of network theory the concept of centrality [Freeman, 1979; Wasserman et al., 1994; Brandes, 2008; Scott, 2000] on tourism research has also highlighted specific characteristics of networked tourism systems.

On of the main applications is to identify who are the most ”important” nodes on the network. Degree centrality is the most intuitive indicator on networks, measuring the number of nodes that an individual is connected to. Freeman [1979] encountered three forms of centrality: degree (see section 3.2.1), betweeness (see section 3.2.3) and closeness. The more central is an individual the more prominent and powerful Brass [1984]; Rowley [1997]; Burt [1995]; Brass and Brurkhardt [1993].

Centrality reveals how an organization is positioned within a global structure of networked organizations [Pavlovich, 2003; Scott et al., 2008b], and how they are related, rather than on the individual attributes [Rowley, 1997]. The more central position an organization has, the so called hubs, the easier it is to access information and it enables a faster action, generating benefic networks Powell et al. [1996]. Peripheral organizations showed to have several benefits on relating with the most ”important” ones, enabling more information and resource accessibility [Baum and Oliver, 1996].

Regarding multi-destination networks show how important is structural relation between destinations within a region [Shih, 2006; Hwang et al., 2006]. It was studied the relation between 16 destinations within a tourism region, by a survey administrated by tourists that have been on the region. The global network was captured on a graph (see section 9.4, Fig. 9.8) and the interdependency of destinations was studied. The centrality measurements showed to be related with resource and traffic, indicating destinations dependency and conductivity. The structural characteristics of a region composed by a set of tourist destinations was depicted by its networked configuration, using centrality indicators and structural holes theory. More over the author shows how destinations form a complementarity of available resources and attractions, resulting on a classification of destinations based on network analysis appropriate for planners and governments for creating new routes or improving the existing one. Shih [2006] concludes that ”on a practical level, increased knowledge regarding the compatibility and complementarity of tourist facilities among multiple destinations can result in more focused of multidestination products.”

A similar studied in Northern Indiana described the tourists behavior [Zach et al., 2008], by using network analysis to study the network structure of tourist attractions. The configuration of the attractions also shows a the the importance of hubs (most central nodes), with a very large number of periphery nodes only connected to the hubs. Further studies on the relational configuration of destinations attractions probably provide a numerous strategic results to better coordinate and plan the destination.

On section 9.4 the multidestination network is explain in more detail and compared to the data analysis of this thesis also concerning structural relations between destinations.

Some more useful concepts of network theory are density and clustering, that related with embeddedness explain collective behavior and stakeholders cohesion on the tourism sector Scott et al. [2008b]; Baht and Milne [2007]. Density, measuring the ration between the number of existing ties over the total number of possible ties, explores the whole network structure.

On interorganization network the notion of density and centrality relates with effectiveness on transfer of institutional norms and stronger information exchange [Meyer and Rowan, 1977; Di-Maggio and Powell, 1983]. Due to be highly connected linkage, stakeholders developed common behavior norms inducing common institutional norms. In this sense, the relation between centrality and density influences the collective configuration of network [Uzzi, 1997; Granovetter, 1973]. Rowley [1997] developed a model on this relation concluding that on less dense networks the central organization has a less prominent role, while on dense networks the central organizations has compromising actions because of their need to conform stakeholders pressure. By other hand in less dense networks the less central organizations the less central organizations tend to have less pressure and behave more in accordance with the central organizations. Therefore, the central organization always play a commander position, called also the hubs of the network.

4.3.3 Complex Networks Analysis

The science of complex networks is complementary to social network analysis on the overall science of network theory. As explained in section 3 the latest developments englobe a significant number of results and several real-world applications ranging from biology to information technologies. Like in general network analysis also complex networks contribute to the understanding of tourism systems.

The latest development on complex networks explain some rules on the evolution of systems or by other words how the relational structure evolves over time. Like in other real-world networks, tourism has showed to depend much on the relational configuration [Scott et al., 2008b; Shih, 2006; Hwang et al., 2006]. Further Hwang et al. [2006] and Shih [2006] conducted studies to show that travel patterns can be understood as networks, by assessing the structural properties of travel within and between different destinations. Additionally Zach et al. [2008] show studied how places visited in Northern Indiana show the effect of ”long tail” (see section 3), relating the finding with the Zipf’s law and the preferential attachment of [Albert and Barabási, 1999].

The increase importance of information technologies is one of the main driving forces changing the way tourism organizations manage and operate [Poon, 1993] and the internet is becoming the main channel to search for traveling information [Buhalis and Main, 1998]. The relations between organizations are close related with their relations within their webpages through hyperlinks, as studied by Baggio [2007]. This study on the island of Elba, coast of Tuscany, Italy, showed the relationships among stakeholders on the web space have statistical properties of a power-law [Baggio, 2007]. This comes along with the general general findings of the structure of the internet configuration [Albert et al., 1999].

The findings on complex networks and travel patterns are a promising field of research to understand the structure and configuration of relationships on tourism systems. Some evolution rules, like preferential attachment (see section 3.3.4), can bring some further knowledge on how competition and cooperation evolves over time. Along with this findings this thesis accomplishes the theory of networks and adds complex networks models to the study of relations between international tourist destinations on a global scale (see section 7).

Part III

Methodology and Network Analysis on Travel and Tourism

5 Methodological Standpoints

Introduction

In this chapter, the methodology of the research is presented. It begins with a presentation of the research path, that is explorative in nature. The chapter outlines the epistemological position of the researcher, see section 5.1. Then, the strategy of the research implementation is sketched. On section 5.3 the explanation of the selection of data collection and analysis process are presented.

To illustrate the network analysis as a useful methodology for studying the structural characteristics of multiple destinations, this study introduces the techniques and indicators of network analysis that are appropriate for examining the structural characteristics of destinations, and then tests these techniques and indicators by examining the network of worldwide tourism destinations.

The objective of this research is to study networked systems in tourism from a structural, interdisciplinary and holistic point of view. In the following section, the epistemological standpoints are described through discussion of the role of a research. The relatively unexplored nature of the phenomenon of tourism networks, a search for the unknown connections between the context of the network, its structure and outcomes calls for a methodological approach where the researcher takes a learning role and achieves new methodologies or, as its believed, that better model and depict tourism as a system, see section 5.2. The undefined nature of networks as a theory and paradigm is discussed on the following section.

5.1 Networks Epistemology and Paradigm

The objective of this section has been to assess the current state of network research and to discuss its role on methodological and theoretical approaches. In the following section, is clarified the epistemological standpoints through discussion of the role of the researcher. The relatively unexplored nature of the phenomenon of tourism networks, a search for the unknown connections between the context of the network, its structure and outcomes calls for a methodological approach where the researcher takes a learning role [Agar, 1996]. Agreeing to the learning role of the researcher produces certain implications. First, the research cannot be heavily dominated by pre-existing hypotheses and theories. This could prevent the researcher from seeing new, unfolding realities. Having said that, the author is far from suggesting the possibility of total objectivity of the research process or pointing to the impersonal nature of the research process.

From an epistemological point of view in tourism research, we look to networks as the manifestation of interconnected agents (or places) processes, that follows the network paradigm urging on sciences. It enables constructing knowledge about the meaning of the interconnected elements. The nature of tourism networks is rather broad, as it brings into a discussion not only how the destinations are important in and among tourism destinations, but also how the natural form of tourism resources are distributed and how organizations interact. As the need for multi approach emerge, the research followed the process of seeking to synthesize broad perspectives, knowledge, skills, interconnections, and epistemology in a learning process.

The interdisciplinary study aimed to facilitate to understand tourism networks, and their coherence, which it is believed from the researcher point of view, they cannot be adequately understood from a single disciplinary perspective. Networks are a fashion topic, as it brings a new perspective over systems, from a holistic stand point, the destination has a network of organizations, the destination as a network of resources, the destination being the place where tourists from different countries interact, exchange ideas and change themselves.

The research came forward from interdisciplinary concept itself as one of the main focuses of study, discussing or criticizing well defined disciplines' and their way of segmenting knowledge on the topic of networks, aiming discussion concepts and interpretations that may be not well defined and achieving for a coherent view of the subject. It was found that cultivating interdisciplinary in just a complex topic, was both possible and essential to analyzing, evaluating, and synthesizing information from multiple sources in order to reasoning networks as a theory for tourism research, and moving from a methodological approach to a paradigm. The shift from the approach of focusing on specialized application of tourism networks (adopting one particular perspective), to the idea of networks and interdisciplinary focuses on the awareness of the whole and the idea of the whole pattern, of form and function as a unity. Interdisciplinary perspective is specially important in Tourism research for two reasons. First because its core knowledge integrates more than one science, and second because traditional science fields are unable or unwilling certain tourism problems.

Therefore the research process conducted brings together expertise from sociology, information systems, management, mathematics and physics. The goal is to set up a theoretical framework to understand networks in tourism and developing a model or way of understanding how local patterns can influence and define global patterns. Some challenges and barriers were encountered on the research process. The concept of interdisciplinary challenges traditional methods of reasoning, imposing the appreciation of different perspectives and methods. The deep and up to date research undertaken follows exact sciences (mathematics and physics), sociology, and the incorporations of this concepts into a general results. The results bring new knowledge on evolution process on tourism systems.

This research aims to demonstrate that (complex) network theory and network analysis [Wasserman et al., 1994; Dorogovtsev and Mendes, 2003] are good tools for investigating the network characteristics of relations in travel and tourism systems. A scientific speciality advances not only in a series of small incremental steps, as hypotheses are proposed, tested, and then revised, but more often moves forward in major jumps and starts [Kuhn, 1970; Price, 1963]. There is an interesting movement of interdisciplinary research on which tourism research ought to be a part, and that can add substantially to our understanding of travel and tourism systems. The goal of this chapter is to reasoning on the state of network theory as a research methodology and theory, through a variety of fields of knowledge. Also aim to assess the emergent paradigm from an interdisciplinary perspective.

Interest in network theory has been rising very quickly for several years now across a wide variety of fields. For example, in physics [Albert and Barabási, 1999; Strogatz, 2001; Newman, 2000; Dorogovtsev and Mendes, 2003], hundreds of researchers have focused on this topic in the last years, much of it due to the Travers and Milgram [1969] small-world research. The latest developments were discussed on section 3. From the considerable amount of results on real world networks a theoretical field of networks emerged. Models and quantifying principles bring strong analysis to depict network structure and evolution rules.

In management consulting networks play a central role on mapping as a standard diagnostic and prescriptive tool [Cross et al., 2000]. As an example of management consulting on the topic of networks see http://www.fas.at/business/en/index.htm. On the research side networks also play a central topic on management, one of the most refereed works sparked by the theory of structural holes [Burt, 1995]. These are just two examples among many. Network theorizing has emerged in virtually every area of organizational inquiry, including leadership [Sparrowe and Liden, 1997], power [Brass, 1984], turnover [Krackhardt and Porter, 1985, 1986], job performance [Mehra et al., 2001; Leavitt, 1951], entrepreneurship [Renzulli et al., 2000], stakeholder relations [Rowley, 1997], innovation [Smith, 1993], profit maximization [Burt, 1995], inter-firm collaboration [Jones et al., 1997], and so on.

The network analysis is called social network analysis (SNA) when applied to a social system. The methods applied on complex networks and SNA are getting into a common framework. Social network analysis is a broad term that incorporates a variety of methods and applications, yielding a research tradition that is beyond the scope of this review to summarize fully. Important threads have included the development of methodologies to characterize networks, including mathematical tools [Wasserman et al., 1994]. Social network analysis focuses on patterns of relations among people, organizations, states, etc [Wellman, 1988; Wasserman et al., 1994]. This research approach has rapidly developed in the past twenty years, principally in sociology, information and communication science and physics. On of the main organizations working on the development of

Social Network Theory is *The International Network for Social Network Analysis* [INSNA, 2008], on a multidisciplinary approach.

Social network analysis has moved from a qualitative metaphor of the concept of networks, converging into an analytic approach, and lately to a paradigm, with its own theoretical statements, methods and research frameworks. The adoption of core network concepts into social science thinking, in parallel developments with other sciences, has perhaps contributed paradoxically to the belief of network theory as a methodology. Although there is a methodological strength, Granovetter [1973] developed the theory of *strength of weak ties* as one of the first steps of the integration of network theorizing.

On economic science, the network technical concepts are also theoretically based, based mostly on the notion of social capital [Portes, 1998], that is clearly a theoretical construct. Some technical notions followed with structural equivalence [Lorrain and White, 1971] and regular equivalence [White and Reitz, 1985] related with the notion of social role. Similarly, the technical notions of clique, n-clique, k-plex (see section 3.2.6). More generally, social capital theory is largely network theory. Embeddedness theory is network theory. Diffusion theory is network theory. Indeed, in subsequent pages we shall argue that almost all of the major perspectives in organizational theory, such as resource dependency and institutional theory, have incorporated or independently invented key elements of network theory. Of course, this discussion begs the question of what is network theory.

Other widely used statistical methods that worked for attribute data do not account to the network context because classical methods assumed independence of observations, which is the ground basis of network data. Performing an analysis of network data was therefore quite daunting, entailing considerable learning of both methods and arcane computer programs, e.g. Batagelj and Mrvar [2002]; Borgatti et al. [2002]. Hence it is natural that methodology becomes a very salient feature of network research.

This first distinction is about methods. Both quantitative and qualitative network approaches take networks as an analytical tool. The quantitative approach, however, considers network analysis as a method of social structure analysis. The relations between actors are analyzed in terms of their cohesion, structural equivalence, spatial representation using quantitative methods such as ascendant hierarchical classification, density tables, block models etc [Wasserman et al., 1994]. The qualitative approach, on the other hand, is more process-oriented. It focuses less on the mere structure of interaction between actors but rather on the content of these interactions using qualitative methods such as in-depth-interviews and content and discourse analysis. Yet, the two methodological approaches are not mutually exclusive but complementary [Sciarini, 1996].

Perhaps the most fundamental characteristic of network theory is the shift from atomistic explanations in terms of attributes of independent cases to the explanation of phenomena in terms of relationships among a system of interdependent actors [Wellman, 1988]. For example, rather than trying to model adoption of innovation solely in terms of characteristics of the adopter (e.g., age and personality type), network theorists posit interpersonal processes in which one person imitates or is influenced by or receives something from another.

This fundamental shift from attributes to relations entails a change in theoretical constructs from monadic variables (attributes of individuals) to dyadic variables (attributes of pairs of individuals) which constitute binary relations among a set of actors. The dyadic ties link up through common nodes to form a field or system of interdependencies we call a network. This gives some network theorizing a holistic or contextualist flavor in which explanations are sought not only within actors but in their network environments, which may include quite distal elements unknown to the actor but linked to them through chains of ties, like the butterfly effect in complexity theory [Lorenz, 1963].

At a more specific level, network theorizing consists of an interplay of the specific properties of ties (i.e., what function they serve) with the topology of ties - the pattern of interconnection. For example, suppose friends within an organization tell each other about the latest office gossip. The supposition is a claim about one of the functions of friendship ties (or the kinds of processes they support). Now, it is reasonable to propose that a person with more ties should receive more news (i.e., have greater probability of hearing any specific item), for obvious reasons [Jones et al., 1997]. This is a bit of network theory, albeit at the simplest possible level.

While we are at it, we can think about whether the strength of ties is independent of the pattern of ties. It seems plausible that if persons A and B share many close friends, they will very likely become at least acquainted, and may be predisposed to like each other. This implies that people are more likely to hear novel information from those they are not close with, since their social circles overlap less [Granovetter, 1973]. The connections to organizational outcome variables such as job performance, mobility and turnover are obvious. It is equally obvious that we can no longer deny the existence of network theory.

In addition, almost all of the hundreds of articles on networks contributed by physicists in the last few years are focused on the evolution of such social networks as the world-wide web, co-authorship among scientists, and collaboration on movie projects. They posit interesting processes of network growth, such as the preferential attachment model in which nodes entering the network preferentially form links with nodes that already have many links, creating a network structure known as ”scale-free” in which the distribution of ties to nodes is not normally distributed but rather follows a power law. A review of this work is provided by Newman [2003a].

One handicap has been the lack of methodological tools and statistical models for modeling network change, but this situation is changing rapidly with the development of new models and computer programs [Batagelj and Mrvar, 2002; Borgatti et al., 2002]. Another crucial development that is likely to spur research on network change is the recent fusion of network research with adaptive agent simulation of organizations.

Social Network Theory comes with its contributions on bringing methods to analyze the networking patterns. Social Sciences have a long history on studying the qualitative and quantitatively real world networks [Freeman, 1979; Scott, 2000]. A network is a set of nodes, also called vertices, with linkage between them, called ties or edges. But networks can be rather more complex. Ties may have a direction and weights. Considering countries as tourist destinations, the countries themselves may be considered as nodes. Nodes may also be organizations, tourists, destinations, theme parks, etc. Ties can also be of various types, depending on what is considered. Some examples are trade, partnerships, tourist flows, cultural proximity, geographic closeness, etc.

The dynamics of the network over time can also be represented, and some other techniques can be used to study the evolution. Sociometry developed on the exploration of the hidden structures of a network, relations, subgroups, cohesion, informal linkage, friendships, different interactions among actors, etc [Wasserman et al., 1994]. In Social Network Analyses questionnaires are mostly used as a data collecting methodology, trying to find the hidden patterns on the relations. The prominence of nodes can be visualized, like centrality and connectivity. If a node is removed how connectivity is affected? How many nodes do we need to remove to affect the network in a certain way? One of the primary tools in networks is the visualization. Being able to have a whole draw of the network brings location and prominence of nodes, more periphery nodes, etc.

New research developments on Network Theory have a new direction, bringing a complementary view. As the networks flourish in the world, the latest methodologies and techniques are promising. Social Network analysts have paid special attention to the study of the patterns of networks with a small number of nodes. The new trend in Network Theory, denominated Complex Network Theory, is to comprehend how networks with large number of nodes evolve. This research has been geared by the availability of computers and data collection capacity based on IT. Nowadays it is possible to manage large databases with information on the relationships of a network. Those can be communication networks, like emails, world-wide-web network, or transportation networks, etc. These large networks involve hundreds or thousands of nodes.

One of the most remarkable characteristics of Network Theory is its rising in diverse sciences, and the constant contribution of each of the sciences to the common body of knowledge. Tourism research has also been related with networking challenges and some conceptual questions arise on the way relations can be studied in tourism. Over the last decade the research was geared towards real world networks, like the Internet, the World Wide Web, business and financial networks, social networks through emails, organizational networks, neural networks, metabolic networks, food chain networks, some others biologic networks, mail post networks, scientific papers citations, among others.

Those studies have been driven mostly by empirical application. Some properties of these networks can be confronted showing that large networks contain a hidden structures. New measures are needed because the precious methods cannot be applied to these networks. For example, if one node from a network with thousands of nodes is removed, there maybe be no difference and that may not even be visualized.

The study of large networks is based on statistical properties that describe the structure and behavior of networked systems. New measurement are developed considering systems requests. The Complex Network Theory seeks to model the system of real networks, and aims to relate the models to the studied properties. Once we would be able to model, a prediction of network system behavior and the clustering and cohesion of local laws could be estimated. Comprehensive and detailed analyzes of the network theory and research are provided by Wasserman et al. [1994]; Scott [2000]; Dorogovtsev and Mendes [2003]; Albert and Barabási [2002]; Watts [2003b]; Buchanan [2003].

5.2 Tourism Sphere

On tourism research, the emergent science of network theory, and its parallel development on methodological and theoretical approach is happening during the last few years. The literature review on the main concepts of tourism networks and its concepts and definitions is discussed on section 4. The main focus is on a qualitative need to define relationships and the agents of the tourism systems. The social structure rising due to globalization, information technologies, knowledge economy, and transportation infrastructures are launching new linkages in society. Tourism is developing in this networked society, and it has also been challenged by the complexity of networking among organizations, tourists, public-private partnerships, etc. Tourism destinations are increasingly competing through networking strategies, requiring an efficient structure at the destinations.

Tourism research is a sphere of knowledge that encompasses methodologies and theories from various other fields of science [Jafari, 2000]. The rapidly grows of tourism, and its impact on country/region/local communities is affecting the research perspective. It has with a truly multi, cross and interdisciplinary sphere of knowledge. This perspective followed all the research path, as a way to locally and globally understand tourism systems. Besides its truly multi-disciplinarian

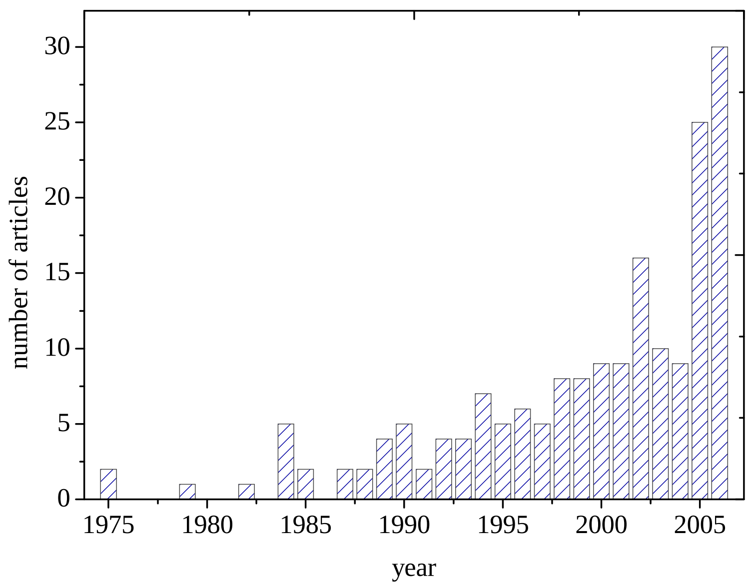


Fig. 5.1: Evolution of the number of papers on the concept of networks. The papers were selected among the journals with higher impact in tourism research [Hall, 2005; Pechlaner et al., 2004; McKercher et al., 2005; Ryan, 2005; McKercher, 2005].

knowledge, tourism is a rapidly growing sector. The growth in global tourism is an attractive opportunity to countries that are willing to compete to the dynamic trends of the global tourism market place.

Attempts have been made to categorize networks on the tourism sector, although nodes and relationships classifications may be highly related with the system that is aimed to represent. The interest of the network concept from a qualitative and quantitative side on tourism research has increased over the years. On Fig. 5.1 is the evolution of the term ”network”, where journals were selected among the ones with higher impact in tourism research [Hall, 2005; Pechlaner et al., 2004; McKercher et al., 2005; Ryan, 2005; McKercher, 2005]. Follows the list of the chosen journals: Annals of tourism, Tourism Management, Tourism Economics, Journal of Travel Research, International Journal of Contemporary Hospitality Management, Journal of Travel & Tourism Marketing, Journal of Leisure Research, Journal of sustainable tourism, International Journal of Hospitality Management and International Journal of Tourism Research.

5.3 Data and Software Programs

On this section are described the type of data collected during the doctoral research. The description of the data is on the section refereed.

- Data Collection
  - Secondary Data
    - Tourist Arrivals: data from the World Tourism Organization, see section 6.
  - Primary Data
    - Webpages on tourism: google search and public and private organizations webpages, see section 9.3.

The research evolved a large amount of mathematical analyzes and computer analysis. The programs used on the different applications were:

- **Python** [McDonald, 1999] was used for the webspace analysis on the places of interest on a given destination, see section 9.3. Python is a dynamic object-oriented programming language. Appropriate for this study for the extensive standard libraries on numerical and statistical analysis.

- **Mathematica** [Wolfram, 2003]: The data analysis of the network measurements, with the theoretical framework presented on section 3.2 and the data analysis on 6, was performed in Mathematica. All the measurements on the data analysis were developed by the researcher.

- **Pajek** [Batagelj and Mrvar, 2002]: The network analysis was reinforced with a visualization tool, Pajek, used to obtain the networks in section 6.2. More specifically Fig. 6.4 and Fig. 6.3. Visualization is a strong tool for network theory, that allows the user to more easily interpret the system relations.

- **UCINET** [Borgatti et al., 2002]: This program is also for representation, analyzes and visualization of nodes (e.g. agents, organizations, or people) and edges (relationships). It was used to verify network measurements of the researcher, on section 6.2.

- **Visual Basic** [Wright, 1996]: The data analyzed of the international tourist arrivals (World Tourism Organization) was on Excel format, as described on section 6. The data was on a separate file for each country, and VB macro on Excel was used to convert it to cvs format.

5.4 Thesis Contributions

Science, culture, socio-economic, politics, etc, are undergoing radical changes over the years and this is driving tourism on a permanent transformation. A global perspective to understand key determinants of market competitiveness is critical for the tourism industry to sustain its growth and vitality. The determinants to have a higher marketshare may be different today and in 2020. This thesis focuses on a theory that enable to understand key relations on tourism systems.

Tourism research is a very challenging field of knowledge, with a truly multi, cross and interdisciplinary sphere of knowledge (see section 5.1), and this perspective followed all the research path, as a way to locally and globally understand tourism in its dimensions. This multi-disciplinarian knowledge is due to the complexity of tourism as an amalgam os services that provide the same product and with its interdependency with a variety of sectors. Besides that, tourism is also a growing sector, rapidly changing due to global technologic and society trends. The growth in global tourism is an attractive opportunity to countries that are willing to compete to the dynamic trends of the global tourism market place.

The thesis has three main contributions, from the point of view of the researcher, that we shall present on this section.

First the conceptual strands on networks in tourism are analyzed as well as its relation with the rich (social) network theory and complex networks theory literature (see section 3). The highly interdependency of tourism stakeholders, resources and destinations are viewed on a common framework which are recognized to follow general results of networks like in other real-world systems.

Therefore a strong contribution of this thesis stands on bringing to tourism research not only empirical evidences on how the configuration and the acting of relations within a network are benefic (see section 4), but more broadly a theorization linked to an old school of ”networkers”, from sociology, management, economics, mathematics and physics. The goal objective of generalization seems rather ambitious, but is based on common empirical evidences on many real-world networks that so naturally can seem disparate phenomenons [Castells, 1996; Barabasi, 2003], and so naturally follow a common structure.

Moreover, a set of quantitative tools brought from network theory are introduced, its application on tourism researched are reviewed. Generally a common framework of indicators can be set up from the used methods along section III. A qualitative versus quantitative approach is also a step forward made by this thesis. The information society not only influenced the new networked age, but also brings new potentialities on the way information is used, collected and analyzed. Therefore, new analytic techniques can be the only way to analyze large sets of data, on many kinds of relations, from informal like emails or travel chatrooms (like TripAdvisor) to formal like interorganizations connections on a destination, able to analyze relations among hundreds or even millions connections between nodes.

Second the results show the first evidences of the human travel evolution network, responsible for about 10% of world's domestic product. Our results thus suggest that being tourism a sector that developed with no specific strategies, the evolution of traveling and tourism networks indeed follow some universal principles (see section 7). These may have quite strong implications on tourism competitiveness and is possible a base knowledge to compete on the international marketplace. Besides that the tourist destinations are disposed on a structural basis that is depicted by a random network (see section 3.3.1), but when considering volumes of tourists or flows the network becomes a scale-free, the also known for a richer gets richer dynamics (see section 3.3.4).

The findings reveal in this way that the tourist destinations network has an homogeneous topology, meaning that countries have approximately the same number of tourist destinations. Even if the arrivals of tourists can be dramatically different, the countries tend to do not diverse much on the number of choices. Therefore, the main rules of human travel are on the volume of the flows rather than on the diversity of connections. The scale-free behavior defines basic competitive dynamics. In general the results are also the first proof that the international human travel network the Zipf law [Zipf, 1949] and the Pareto distribution [Pareto, 1896].

The network of tourist destinations is also shown to be dissassortative (see section 3.2.5 and 8.1). Rather than having social backbones where the central destinations connect to other central destinations, called hubs, the overall structure behaves on connecting hubs to periphery and otherwise disconnected destinations. This structural characteristic is typical from economic and technologic networks, and not from social networks. Moreover the dissassortative behavior is not only obtained on the world tourism network, but by comparing with an interorganizational network it is proposed that a dissassortativity can be a general property on the travel and tourism industry (see section 9).

Third, on some tourism and non-tourism real-world networks are observe common behaviors. While the random versus scale-free structure of the world tourism network is similar to a regional urban traffic network, the scaling evolution also approximate the international airports network. Although the airports network is assortative, showing a strong political influences on its structure, which is dissimilar to the tourism networks that lays around 40% of its flows on the air transportation. The researcher proposes in this way that this dissimilarity and regional politics influencing airports structure can be influencing the demand for charters (see sectionn 9.2).

Future research on tourism networks is needed to accomplish more results on the theorization of relationships in tourism, and the way they involve over time. Better linkages can dramatically influence competitiveness of tourist destinations. Therefore, the researcher focused on analyzing the different methodologies and results from different knowledge areas (see section 3), discussed their applications to tourism research to develop a set of indicators. For tourism sector the relationships play a special important role for destinations, due to the interdependency between organizations and destinations.

International Tourism Network 6

Introduction

On this section is introduced the data, as well as its description and discussion, comprimising the world tourist arrivals and departures between every two country in the world, from the UNWTO. The data is analysed on the following sections.

Tourism industry is a group of economic activities which combined makes it the world's largest industry, the number one generator of jobs, one of the world's biggest exports, and a major stimulus for investment and growth, having a positive effect on many countries through out the world. It is believed that tourism will continue to be one of the most dynamic growth sectors of the global economy.

For decades international tourist arrivals have highly increased. On the year of 2007 were registered 898 million international tourist arrivals (on 208 countries and territories). Further WTO expects that number to reach 1.6 billion by 2020 WTO [2000].

Having tourism as one of the fastest growing economic sectors, questions arise on how is the network of tourists/travelers evolving? Where are these tourists going and coming from? Which are the motivation for new traveling connections? What is the relationship among countries that are popular or unpopular? The research aims to answer some of these questions, and to develop some methodological tools analyzes travelers behavior. Besides revealing information on human traveling patterns, the network will also likely help future research understand information transfer and global wealth flows, since tourism accounts for more than a significant part of the world's gross domestic product.

6.1 Worldwide Tourist Arrivals: Data Description

The research data regarding the world tourism network was based on secondary data from the UNWTO (United Nations and World Tourism Organization). Quantitative tourism-related data presented is based on a selection of data included in the UNWTO database on World Tourism Statistics WTO [2006]. This data collection from the WTO contains a variety of series for over 208 countries and territories covering data for most countries from the 1980's on and is continuous updated. The methodology and data selected to our research is discussed and analyzed in this section.

The statistical data has been collected by the UNWTO Secretariat from the official institutions of the countries and territories (UNWTO member as well as non-member countries) or from official international bodies, e.g. the Caribbean Tourism Organization (CTO), International Monetary Fund (IMF). The data for individual countries are based on full year results, or projections, as communicated to the UNWTO Secretariat by the authorities of the countries and territories or disseminated through a news release, publication or on the Internet.

In the world and (sub)regional aggregates, estimates are included for countries and territories with data still missing based upon data available for a part of the year or the general trend for the region. In particular for the Middle East and Africa the regional and subregional aggregates for 2002 should be treated with caution as estimations are based on a relatively small number of countries and territories that supplied data for the entire year.

UNWTO tourism statistics generally refer to figures for a country as a whole. In the collection of statistics, however, except for independent states, there are also a number of dependencies or territories of special sovereignty included (for instance Hong Kong (China) or French Polynesia). These territories report tourism figures independently and are for the sake of tourism statistics considered as an entity in itself. Because of this, where reference is made to ”countries” the term generally should be taken to mean ”countries and territories”. In a few other cases, dependencies are not separately listed but included in the total for the country they depend upon (for instance Guernsey, Jersey and the Isle of Man in United Kingdom).

In general UNWTO does not collect data on the level of regions, states, provinces or specific destinations within a country (Hawaii is one of the few exceptions made because of its relevance for Asian outbound travel; in the overview tables, however, Hawaii is included in the United States figure).

Despite the considerable progress made in recent decades, international tourism definitions and methods of data collection tend to differ. Regarding an international comparability of statistical data more general methods should be chosen. The methodology of UNWTO on collecting tourism statistical data has taken considerable changes on the beginning of the ninetinth decade, for this reason it was not recommended by the UNWTO to use older data. The Benford law was calculated to test the reliability of the data on section 6.3.

The regional country groupings are according to the UNWTO regional and subregional grouping, World: Africa, Americas, Asia and the Pacific Europe and Middle East, as listed below.

- **Africa**
  - **North Africa:** Algeria, Morocco, Sudan, Tunisia
  - **West Africa:** Benin, Burkina Faso, Cape Verde, Côte d’Ivoire, Gambia, Ghana, Guinea, Guinea-Bissau, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, Togo
  - **Central Africa:** Angola, Cameroon, Central African Republic, Chad, Congo, Democratic Republic of Congo, Equatorial Guinea, Gabon, Sao Tomé e Princípe
  - **East Africa:** Burundi, Comoros, Djibouti, Eritrea, Etiopía, Kenya, Madagascar, Malawi, Mauritius, Mozambique, Reunion, Rwanda, Seychelles, Tanzania, Uganda, Zambia, Zimbabwe
  - **Southern Africa:** Botswana, Lesotho, Namibia, South Africa, Swaziland

- **Americas**
  - **North America:** Canada, Mexico, United States
  - **Caribbean:** Anguilla, Antigua Barbuda, Aruba, Bahamas, Barbados, Bermuda, Bonaire, Bristish Virgin Islands, Cayman Islands, Cuba, Curaçao, Dominica, Dominican Republic, Grenada, Guadeloupe, Haiti, Jamaica, Martinique, Montserrat, Puerto Rico, Saba, Saint Lucia, St.Eustatius, St.Kitts-Nevis, St.Maarten, St.Vincent,Grenadines, Trinidad Tobago, Turks and Caicos, US Virgin Islands
  - **Central America:** Belize, Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, Panama
  - **South America:** Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, French Guyana, Guyana, Paraguay, Peru, Suriname, Uruguay, Venezuela
- **Asia and the Pacific**
  - **North-East Asia:** China, Hong Kong (China), Japan, Democratic people's republic of Korea, Republic of Korea, Macao (China), Mongolia, Taiwan (pr. of China)
  - **South-East Asia:** Brunei Darussalam, Cambodia, Indonesia, Lao P.D.R., Malaysia, Myanmar, Philippines, Singapore, Thailand, Vietnam
  - **South Asia:** Afghanistan, Bangladesh, Bhutan, India, Iran, Islamic Republic of, Maldives, Nepal, Pakistan, Sri Lanka
  - **Oceania:** American Samoa, Australia, Cook Islands, Fiji, French Polynesia, Guam, Kiribati, Marshall Islands, Micronesia (Fed.St.of), North Mariana Islands, New Caledonia, New Zealand, Niue, Palau, Papua New Guinea, Samoa, Solomon Islands, Tonga, Tuvalu, Vanuatu
- **Europe**
  - **Northern Europe:** Denmark, Finland, Iceland, Ireland, Norway, Sweden, United Kingdom

- **Western Europe:** Austria, Belgium, France, Germany, Liechtenstein, Luxembourg, Monaco, Netherlands, Switzerland
- **Central and Eastern Europe:** Armenia, Azerbaijan, Belarus, Bulgaria, Czech Republic, Estonia, Former U.S.S.R., Georgia, Hungary, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Poland, Rep Moldova, Romania, Russian Federation, Slovakia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan
- **Southern Europe:** Albania, Andorra, Bosnia Herzgovina, Croatia, F.Yug.Rp.Macedonia, Greece, Italy, Malta, Portugal, San Marino, Serbia Montenegro, Slovenia, Spain
- **East Mediterranean Europe:** Cyprus, Israel, Turkey

- **Middle East**
  - Bahrain, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libyan Arab Jamahiriya, Oman, Palestine, Qatar, Saudi Arabia, Syrian Arab Republic, United Arab Emirates, Yemen

Defining tourism can be a difficult task, as it accomplishes a complex product, dependeing on the interdependency of different organizations and competing on a constantly growing compatitive international market. As described on section 2, the definition of tourism evolved over the centuries. To our consideration we follow the definition of the world tourism organization.

Definition 6.1 ***Tourism:*** *According to the UNWTO/United Nations Recommendations on Tourism Statistics, tourism comprises the activities of persons traveling to and staying in places outside their usual environment for not more then one consecutive year for leisure, business and other purposes.*

This concept can be subdivided, depending upon whether a person is traveling to, from or within a certain country the following forms can be distinguished. For definition of domestic tourism see Def. 6.1.

Definition 6.2 ***Inbound Tourism:*** *Involving the non-residents received by a destination country from the point of view of that destination.*

Definition 6.3 ***Outbound Tourism:*** *Involving residents traveling to another country from the point of view of the country of origin.*

For example, portuguese tourists traveling to Spain are accounting as a part of the inbound tourism to Spain and as part of the outbound tourism of Portugal. Each of the countries provides information about tourist arrivals from each of the other countries, or each of the countries provides the tourist arrivals of their inbound tourism. On Tab. 6.1 is an example of the data provided by Spain for the years $1999 - 2001$. The countries are separated by region and subregion, have a number and a ISO3 name.

This infromation was provided on EXCEL files, by the UNWTO, so for the research it is secondary data. It was provided a file for each country. A macro on Visual Basic was created to read all the files at onces and extract the information to a .cvs format with the information on a readable format for programming languages. The .cvs files were posteriorly analysed and converted to an adjacent matrix, see also section 6.2.

All types of travelers engaged in tourism are described as visitors. Visitors can be distinguished as same-day visitors or tourists (overnight visitors). There are various units of measure to quantify the volume of tourism, as described below:

- **Visitors**: measured on the arrivals, at frontiers, or at a specific place in case of domestic tourism.
- **Tourists** (overnight visitors):
  - **Arrivals**
    - at frontiers
    - at hotels and similar establishments [1]
    - at collective tourism establishments (e.g. hotels and other)[1]
  - **Nights** at hotels and similar establishments

[1]Excludes tourism in private accommodation; Arrivals are counted in every new accommodation visited.

Tab. 6.1: Sample of the data for the year $1999 - 2001$, Spain. Arrivals of non-resident tourists at national borders, by nationality. Source: [WTO, 2006].

ESP	**724**	**Spain**			
REG.	**COD.**		**1999**	**2000**	**2001**
		Total	46.775.870	47.897.915	50.093.557
2	**20000**	**Americas**	**2.518.590**	**2.174.344**	**2.079.643**
2	**23000**	**North Amer**	**1.317.958**	**1.491.853**	**1.453.626**
23	124	Canada	104.709	118.417	144.412
23	484	Mexico	179.149	227.807	172.947
23	840	USA	1.034.100	1.145.629	1.136.267
2	**24000**	**South Amer**	**648.166**	**629.320**	**546.851**
24	32	Argentina	266.169	242.367	178.803
24	76	Brazil	219.173	203.015	164.592
24	152	Chile	74.206	91.877	66.117
24	862	Venezuela	88.618	92.061	137.339
2	**25000**	**Other America**	**272.024**	**397.417**	**173.867**
25	**932**	**Other America**	**272.024**	**397.417**	**173.867**
3	**30000**	**East As/Paci**	**359.113**	**300.828**	**265.047**
3	**31000**	**N/East Asia**	**359.113**	**300.828**	**265.047**
31	392	Japan	359.113	300.828	265.047
4	**40000**	**Europe**	**43.587.461**	**44.499.469**	**46.827.387**
4	**41000**	**C/E Europe**	**212.544**	**237.459**	**298.863**
41	643	Russian Fed	212.544	237.459	298.863
4	**42000**	**Northern Eur**	**14.886.199**	**16.130.400**	**17.840.693**
42	208	Denmark	629.105	613.984	670.401
42	246	Finland	423.873	431.271	436.569
...	...	...	...	...	...

The measure were the following:

1. Arrivals of non-resident tourists
   (a) at national borders
      i. by nationality
      ii. by country of residence
   (b) in all types of accommodation establishments
      i. by nationality
      ii. by country of residence
   (c) in hotels and similar establishments
      i. by nationality
      ii. by country of residence
2. Arrivals of non-resident
   (a) visitors at national borders
      i. by country of residence
      ii. by country of residence
3. Overnight stays of non-resident tourists
   (a) in all types of accommodation establishments
      i. by nationality
      ii. by country of residence
   (b) in hotels and similar establishments
      i. by nationality
      ii. by country of residence

Besides the information for each country on the tourist arrivals, the data also encompasis more detailed information for the inbound tourism. For example Portugal (see Tab. 6.3), the arrivals are divided into visitors, tourists, some-day visitors and cruise passengers. It allows to know more about the stay duration of the visitors and the importance of each category over all the tourists.

Considering the regions where the tourists come from the data is also grouped by the sixth regions: africa, americas, europe, east asia and the pacific, south asia and middle east. Although this division can be extracted by Tab. 6.1, some estimeted data can be included for each category, slightly influencing the final results.

On section 2 was mentioned the importance of each transportation means for the world tourism (see Fig. 2.3). The significance of each transport mean is also included on each country data, see Tab. 6.3. For example for Portugal on the year of 2003 the number of inbound tourists arriving by road was significantly higher than the ones arriving by air, around four times more, while rail and sea have quite low percentages. The high number of tourists arriving by road can be related with the fact that around 40% of the tourists are spanish, and therefore, due to the short distance road is the cheaper and in some cases faster mean of transportation. The low-cost airports being constructued in Portugal and low-cost airline companies, along with the fast train connection between Madrid and Lisbon and other socio-economic growth can significantly influence the means of transportation chosen by the tourists.

Other information on the data is the purpose of the visit, divided into leisure, recreation and holidays and business and professional. Also the accommodation is divided into: overnight stays in hotels and similar establishment, guests in hotels and similar establishments, overnight stays over all types of accommodation establishments and average length of stay of non-resident tourists in all accommodation establishments (measured by nights). The total expenditure in the country is considered within the total travel and also by the passenger transport.

In most European countries, domestic tourism is more important for the industry than international tourism. For example for Portugal 58% of all tourists. The domestic tourism is defined as:

Tab. 6.2: Data sample for Portugal, indicators on inbound tourism, year 2003. Source: [WTO, 2006].

Portugal	Inbound Tourism	Units	2003
	Arrivals		
1.1	Visitors	('000)	27.532
1.2	Tourists (overnight visitors)	('000)	11.707
1.3	Same-day visitors	('000)	15.535
1.4	Cruise passengers	('000)	290
	Arrivals by region		
2.1	Africa	('000)	481
2.2	Americas	('000)	10.888
2.3	Europe	('000)	40
2.4	East Asia and the Pacific	('000)	...
2.5	South Asia	('000)	...
2.6	Middle East	('000)	...
	Arrivals by means of transport used	('000)	
3.1	Air	('000)	5.162
3.2	Rail	('000)	92
3.3	Road	('000)	21.961
3.4	Sea	('000)	317
	Arrivals by purpose of visit	('000)	
4.1	Leisure, recreation and holidays	('000)	...
4.2	Business and professional	('000)	...
	Accommodation	('000)	...
5.1	Overnight stays in hotels and similar establ.	('000)	22.714
5.2	Guests in hotels and similar establ.	('000)	4.824
5.3	Overnight stays, all types of accommod. establ.	('000)	24.369
5.4	Average length of stay of non-resident tourists		
	in all accommod. establ.	Nights	4,60
6.1	**Tourism expenditure in the country**	US$ Mn	7.886
6.2	Travel (*)	US$ Mn	6.937
6.3	Passenger transport (*)	US $ Mn	949

Tab. 6.3: Sample of the data analyzed. The data is from $1999 - 2004$ and our sample from year 2002. Arrivals of non-resident tourists at national borders, by nationality. Source: [WTO, 2006]

	DOMESTIC TOURISM		
	Accommodation		
7.1	Overnight stays in hotels and similar establ.	('000)	10.404
7.2	Guests in hotels and similar establ.	('000)	4.714
7.3	Overnight stays in all types of accommod. establ.	('000)	16.457
7.4	Average length of stay of resid. tour. in all accommod. establ.	Nights	

Definition 6.4 ***Domestic Tourism:*** *Involving residents of a given country traveling within that country.*

The tourist that take holidays on their own country, or domestic tourism, account on the data provided by the UNWTO for overnight stays in hotels and similar establishments, guests in hotels and similar establishments, overnight stays in all types of accommodation establishments and average length of stay of residence tourism in all accommodation establishments.

6.2 Network Description

It is analyzed data gathered by World Tourism Organization on the year 2004 WTO [2006] to build a network with 206 countries and establish links - 10886 - between pairs of countries, representing 763 million tourist arrivals, where data for individual countries are based on full year results. It allows a holistic perspective of the worldwide tourism system, rather than a local analysis. The data is 3 years old, although the resulting worldwide tourism network is in effect indistinguishable from the network obtain if using data collected in 2005 (most recent data available).

It was use the data gathered by WTO over these 208 in which countries and territories are considered *nodes*, N, and an *edge* exists from node i to node j when there are tourists from country i to country j. The nodes in the network are the countries and territories while the links (edges) show relationships or flows between the nodes. The network approach is providing both a visual and a mathematical analysis of human traveling patterns.

Fig. 6.1: Representation of the weighted and directed structure, the tourist arrivals between Portugal and Spain on the year 2004.

This network is considerable directed, as for example, the number of tourists traveling from Portugal to Spain is much smaller than the number of tourists traveling from Spain to Portugal. The edge from i to j is different from the edge from j to i, respectively $i \rightarrow j$ and $j \rightarrow i$. On our case we have 5775 edges, L, – representing arrivals of tourists from one country to another, on the year of 2004.

On a directed network the nodes have *in* and *outdegree*, where the *in-degree* of a node i, $k_{in}(i)$, is the number of nodes directed to node i. For example, the *in-degree* of Portugal is the number of countries from which Portugal receives tourists. We may interpret it as the *diversity of inbound tourism*. The *out-degree* of i, $k_{out}(i)$, is the number of nodes that i is directed to. One can also interpret it has the *diversity of outbound tourism*, the number of countries to which Portugal has tourists going to.

The *in-degree* of a country is the number of other countries to which it has national tourists traveling to, in tourism terms representing the number of countries of the *inbound* tourism. The *out-degree* of a country is the number of other countries from which it has tourists traveling from, in tourism terms representing the number of countries of the *outbound* tourism. This type of matrix describes what network analysis calls sociometric choices, which merely depict the presence or absence of a given type of relation [Degenne and Forse, 1999].

Then, summing up the matrix of every tourist yields a valued matrix, in which the (i, j)th cell carries a number that expresses the number of times the tourist routes occur from destination i to destination j. The numbers in the adjacency matrix indicate weightings.

The weighed analysis of the worldwide tourism network is essential because of weights heterogeneity. The network can be expressed by its adjacency matrix $A = \{a_{ij}\}$, dimension N $\times$ N , where $a_{ij} = 1$ if and only if there is an edge from i to j, and $a_{ij} = 0$ otherwise. The weighted adjacency matrix is $W = \{w_{ij}\}$, where w_{ij} equals the flow from i to j. Notice that w_{ij} represents the weight of the edge $i \rightarrow j$ and w_{ji} represents the weight of the edge $j \rightarrow i$, so w_{ij} and w_{ji} are different.

The present network is asymmetric and weighted. The range of the weights goes from 0 to $19.369.677$ with an average value of 81.813, revealing a high heterogeneity of weights. See figure 6.4. The highest value, $19.369.677$ accounts for the tourists from United States of America travelling to Mexico.

As an example on Fig. 6.2 is shown the outbound tourism for Germany in 2004. Each

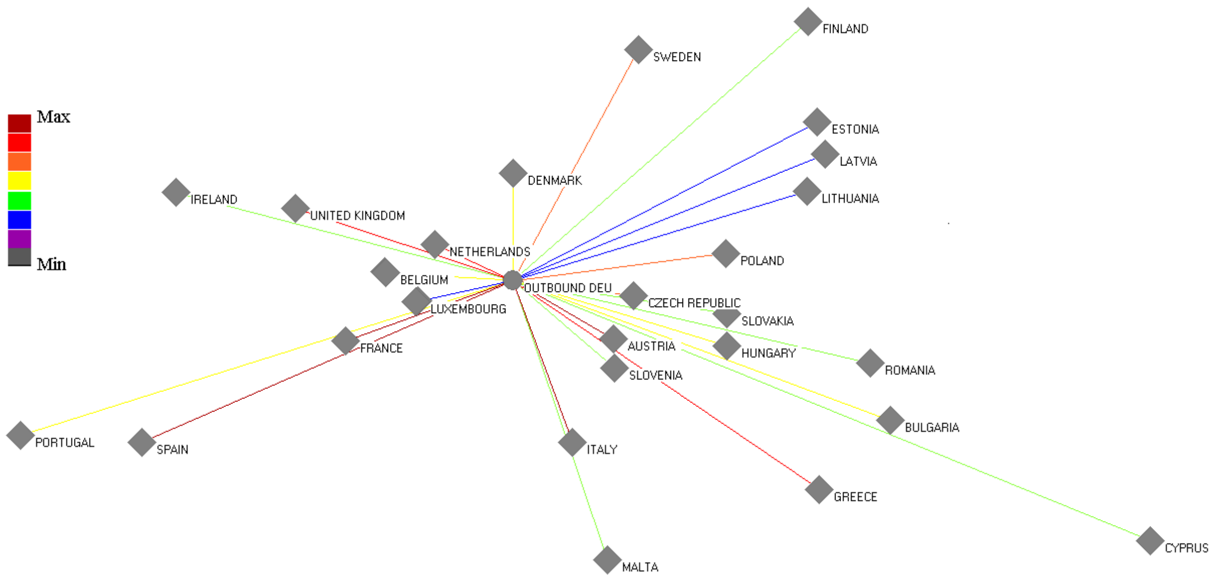


Fig. 6.2: Outbound tourism of Germany in Europe, in 2004. Each lines represents a flow of tourists from Germany to the other european countries. The colors represent the volume to the flow, tourist dapartures, and are on a logarithmic scale.

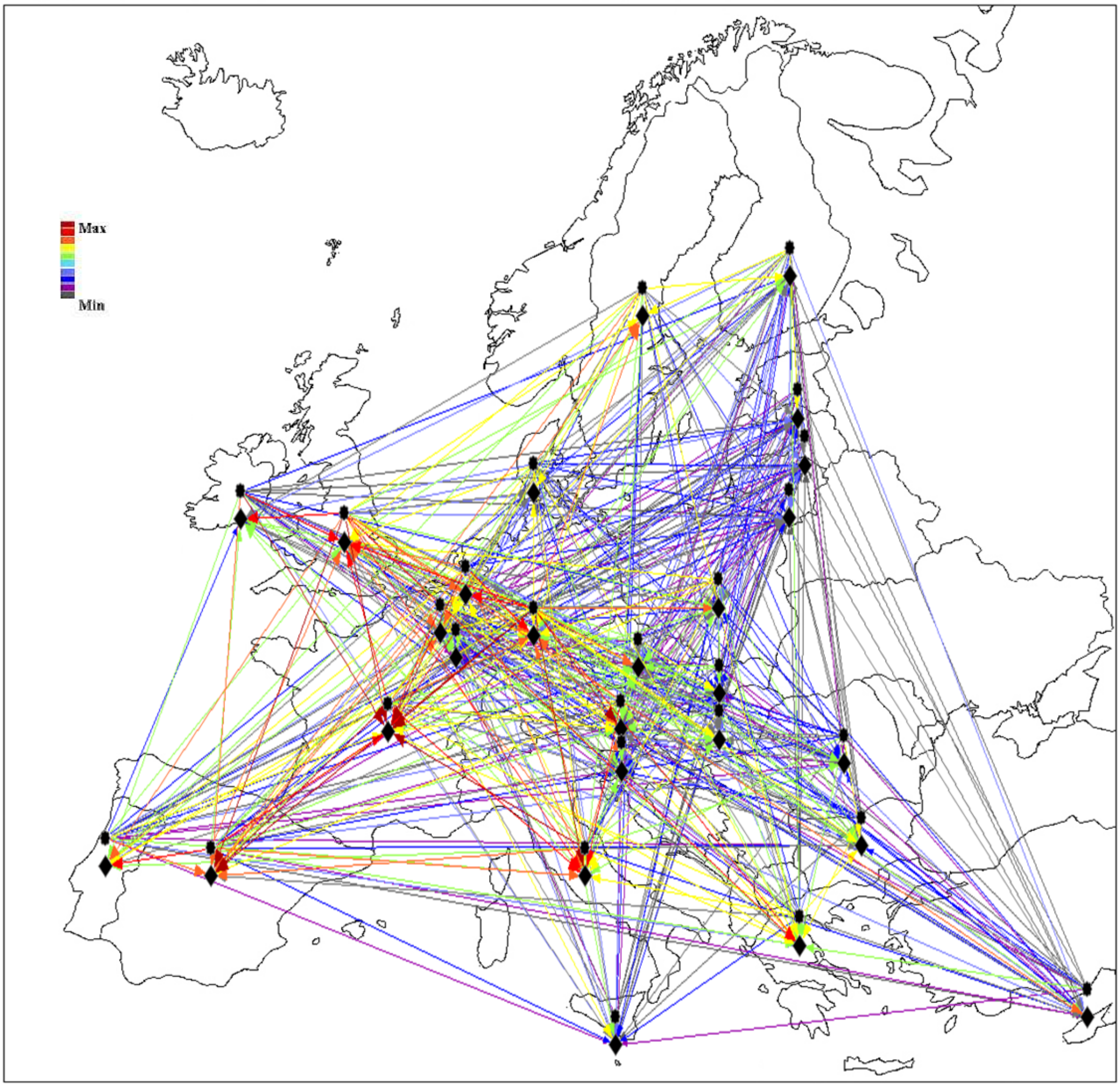


Fig. 6.3: Tourism network of the European Union tourist arrivals. Representation of the weighted and directed structure, on the year 2004. Each lines represents a flow of tourists between two countries.

country has a inbound and outbound tourism that mapped into a network of tourist arrivals in europe we end up with the map on Fig. 6.3.

The network representing the international tourism network accounts for the 208 countries and territories and represents 763 million tourist arrivals, see Fig. 6.4. With 10886 links it is the data representing the highest human travel ever.

International tourist arrivals are analyzed to study *inbound* tourism and *outbound* tourism. *Inbound* tourism, involving the non-residents received by a destination country from the point of view of that destination. *Outbound* tourism, involving residents traveling to another country from the po of view of the country of origin.

The inbound and outbound tourism of a country, in the general international network have a considerable relation, although to some of the countries it is significantly different. On Fig. 6.5 the outbound tourism that in network terms is named strength out, and the inbound tourism, named strength in, have a (pearson) correlation of 0.66, supporting the point of view that the network is considerable directed, also for the strength. For example, Spain had an inbound tourism, strength

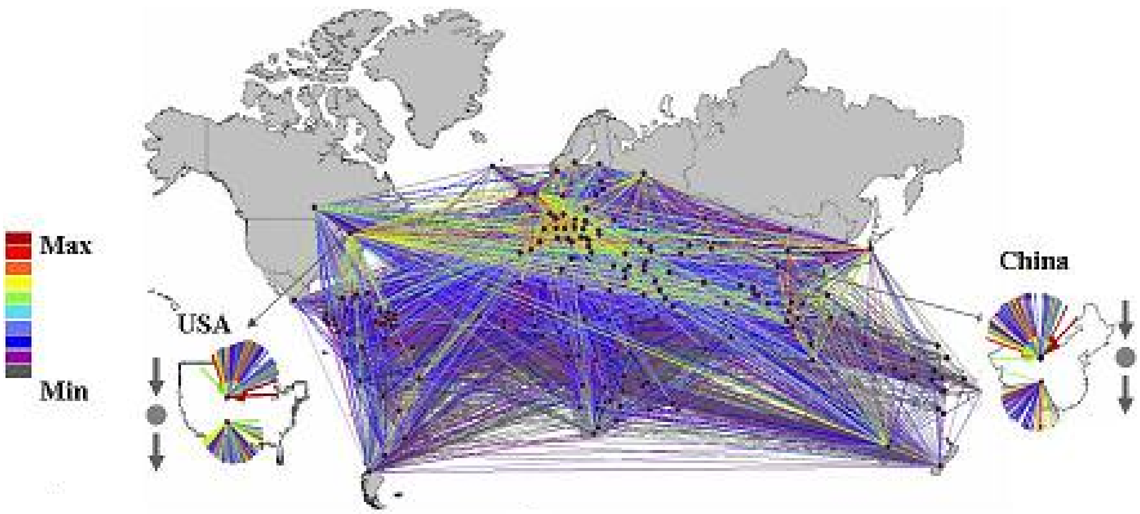


Fig. 6.4: Tourism network on the world map. Representation of the weighted and undirected structure. Each lines represents a flow of tourists between two connected countries. Notice that the higher edge is between USA and Mexico, having a thicker line.

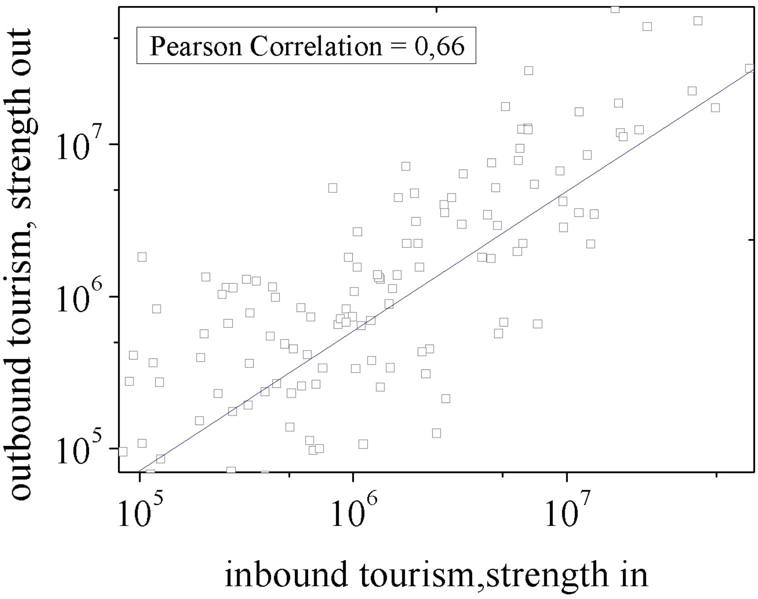


Fig. 6.5: The tourism network is considerable directed, also inbound and outbound tourism have a highly distinct value.

in, of around 50 million tourists in 2004 and regarding outbound tourism, or strength out, around 17 million tourists.

Regarding the domestic tourism, no relation with inbound tourism and outbound tourism was found, both with pearson correlations lower than 0.4. Although the relation between the degree out and the domestic tourism with pearson corelation of 0.69 may tell us about how the variety of tourist destination can be related with the domestic tourism. Notice that with a relation of domestic tourism $= k_{out}^2$ and number of tourit destination have a power-law fit.

6.3 Benford Law in Tourism

On this section is presented the Benford law applied to the tourist arrivals data. The reliability of the world arrivals per country can be regarded as being low, even if all studies on country destination

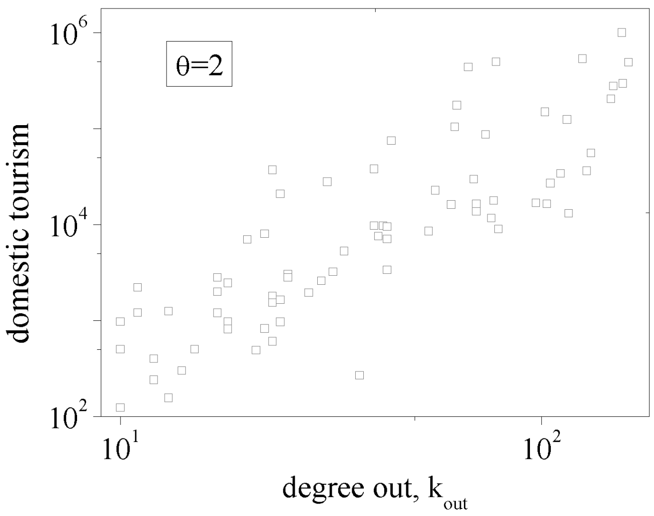


Fig. 6.6: Domestic tourism and tourist destination connectivity, a log-log plot of domestic tourism over k_{out}, domestic tourism $\approx k_{out}^2$.

ranking and forecasting by world organizations is based on it. Here is introduce a method that is commonly used for detection of frauds and mistakes, mostly on finantial issues. Although it is worth noticing that the Benford law, even if having empirical evidences on many case studies, it has no universal explanation. This method is used on the tourist arrivals to also test the data reliability.

Benford [1938] studied at a large amount of naturally occurring data (like the tables of molecular weights, population sizes, river basin drainage areas, and numbers appearing on newspaper front pages) and showed that the leading digits tend not to be uniformly distributed.

For instance, numbers tend to start with the digit 1 in over 30% of his data, but with the digit 9 in less than 5% of the data. To model such data, he proposed an (empirical) law which predicts that the leading digit,

$$P(d = i) = log_{10}\left(1 + \frac{1}{i}\right). \tag{6.1}$$

Benford's Law can be traced back to Newcomb [Newcomb, 1881] who observed that tables of logarithms were dirtier at the front than the back, and proposed without explanation the equation 6.1 for predicting the frequency of the leading digit. The law was then largely forgotten until [Benford, 1938]. This law has been observed on the frequency of occurance of numbers in different real-world systems [Pietronero et al., 2001]. Recently, Nigrini has shown that many aspects of financial accounts like expenses claims follow Benfords Law [Nigrini, 1996]. Using standard statistical tests, he is able to detect fraudulent or erroneous data which deviates from the law.

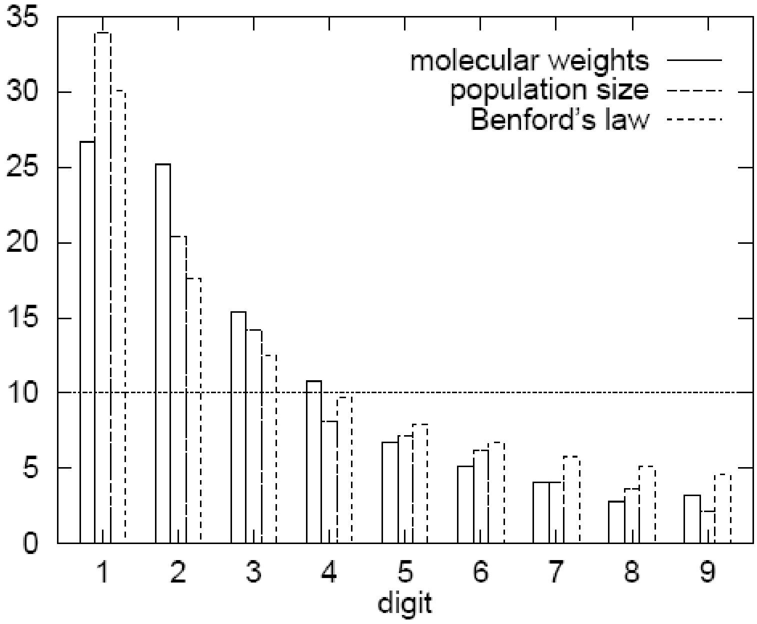


Fig. 6.7: Benford law on the distribution of molecular weights and population sizes. Source: [Benford, 1938]

Many systems have been empirically studied, in economics, management, demographics, etc, which over the years show similar patterns. Whether studying size of cities, people's income or tourist arrivals, is exhibited a power-law behavior. In accordance with this studies, other real-world systems appear to have the same behavior, like number of telephone calls by each user, websites "hits", earthquakes size, etc.

Originally Benford showed the distribution of the leading digits for molecular weights with a sample of $1,800$ different molecular weights and for the ”population size” with a sample of $3,259$ different population sizes, see Fig. 6.7.

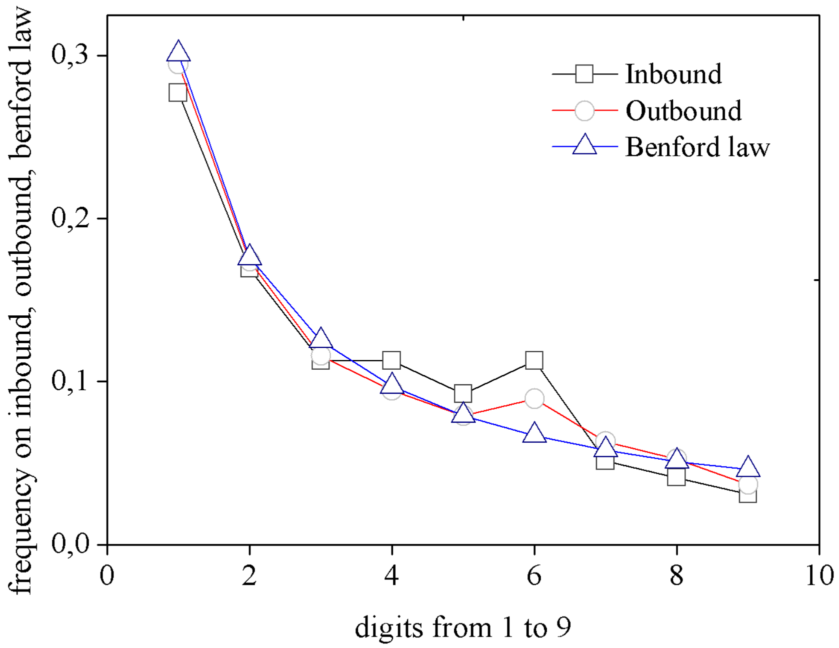


Fig. 6.8: Benford law on worldwide tourist arrivals. Distribution of the leading digit of inbound and outbound tourism.

We show here that Benford's Law models not just natural phenomena like population sizes but also on traveling phenomena. These results could be used to improve how we model such statistics. In addition, by using Benford's Law, we can test generate benchmark sets that may be more realistic than purely uniform random data.

Follows the algorithm to test the benford law:

1. create two lists, e.g. named arrivalsListIn and arrivalsListOut, respectively with data of the inbound tourism and outbound tourism per each country;

2. insert into two lists, e.g. called leadingDigitIn and leadingDigitOut, the leading digit of each of the elements of list arrivalsListIn and arrivalsListOut;

3. plot the distribution of leading digits list;

4. plot in the same graph the benford distribution.

Besides a way of testing the reliability of the data the benford law on the traveling flows can be more realistic to generate data on traveling flows.

6.4 Local Correlations and Transitivity

In this research we use a network approach to study international tourism on the year of 2004. International tourist arrivals reached a record of 763 million in 2004. The international arrival of tourist is yearly measured by the World Tourism Organization (WTO, the major intergovernmental body concerned with tourism) over 208 countries and territories around the world [WTO, 2006]. Worldwide earnings on international tourism reached in 2004 a new record value of US 623 billion.

A very important theoretical idea, reciprocity, was studied and evaluated from the beginnings of social network analysis in the 1930's. The question, first asked about relations such as affected, is, How strong is the tendency for one actor to choose another, if the second actor chooses the first? The several indices of mutuality had also an important influence on the study of reciprocity. We will use the term dyad when specifically referring to a group of exactly interaction of two nodes. We focus on the discussion of the *dyad census*, the counts of the different types of dyads (assuming that specific distributions are appropriate), and tests for hypothesi about the number of choices and the number of mutual choices on a specific relation.

Dyads and triads refer to a group of exactly interacting two and three nodes. The structural properties emerging from this local interactions can show specific tendencies of transitivity and reciprocity. The study the dyads we analyze all the possible dyads, **M** mutual, **A** asymmetric, and **N** null. The triple $\langle M, A, N\rangle$. To study the dyads on the worldwide tourism network we analyze all the possible dyads, one-to-one relations. These could be a mutual (M) relation, $(j \to i) \wedge (i \to j)$, asymmetric (A) relation, $(j \to i) \vee (i \to j)$ (exclusive $\vee$), or a null (N) relation, neither $j \to i$ or $i \to j$.

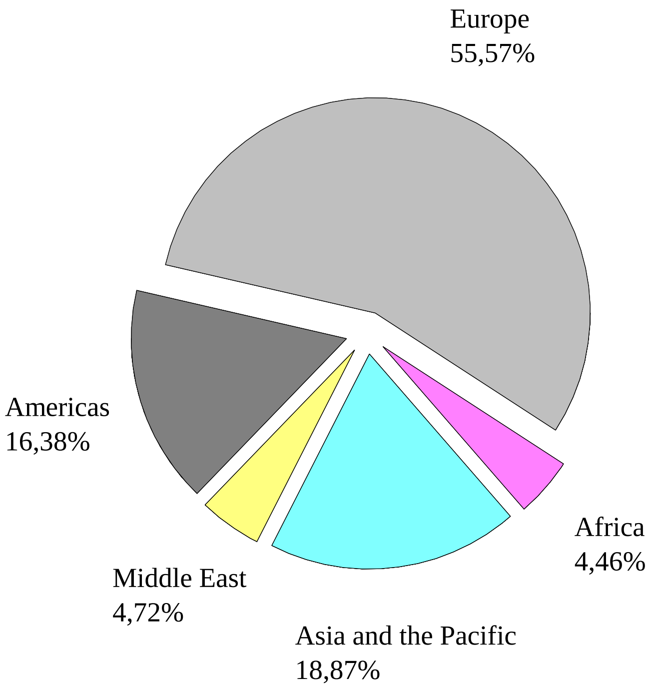


Fig. 6.9: Market share (%) of the international tourist arrivals by region.

Tab. 6.4: Analysis of dyads and triads: reciprocity, density and transitivity; on the year of 2004.

	Reciprocity %	Density %	Transitivity %	Tourist Arrivals
World	27	25	66	765.1
Europe	44	59	85	424.4
America	40	41	75	125.7
Middle East	53	42	69	36.3
Asia and the Pacific	55	64	89	144.2
Africa	21	24	71	34.5

In our data we find that $M = 926$, $A = 5109$, and $N = 15080$. Which means that we have a very small fraction of mutual dyads ($926/21115 = 4,4\%$), considerable fraction of asymmetric dyads ($24,2\%$), and a higher fraction of nulls ($71,4\%$). Between the countries that share tourists we have $15,3\%$ of mutual dyads and $84,7\%$ of asymmetric dyads. See table 6.4. A dyadic analysis seeks to answer several questions about these three values.

An important statistical property to directed networks is reciprocity Wasserman et al. [1994], meaning on the tourism network the appetency to exchange tourists. The links in the network are composed by 4,4% bidirectional links and 24,2% of asymmetric links. If country *j* has tourist arrivals from country *i*, then the probability that country *i* has tourist arrivals from *j* is only about $\frac{1}{4}$, so the network is significantly directed. Notice also that 71,4% of all the pairs of countries are not connected to one another, many new connection can still emerge.

The tourism regions defined by the WTO (World Tourism Organization) are: Europe, America, Middle East, Asia and the Pacific, and Africa. On Tab. 6.4, the reciprocity is measured among countries of each of this five regions, and compared with the whole world. It shows that the reciprocity is much lower in a world scale, than in each of the regions. The exception of Africa may be related with its geographic characteristic of large continent and few amount of tourists. The appetency to exchange tourists is higher on a regional scale, showing that local interaction do play an important factor on flow dynamics (see also Fig. 6.9). Reinforcing this statement, the density of the network is also higher on a regional scale, than in a global scale, which means that tourism movements and interactions tend from local to global.

The observed for dyads, reciprocity and density, is also noted on a three nodes relation, triads. The propensity to have a transitive [Karlberg, 1997] relation is the property that considers patterns of triples of actors in a networks. A relation is transitive if $i \rightarrow j$ and $j \rightarrow k$, then $i \rightarrow k$. Transitivity is the basic relation of triads, and the first measurement applied to a three nodes relationship.

6.5 Conclusions

The movement of tourists on a worldwide scale is responsible for a traveling mobility of hundred millions tourist arrivals every year, representing the largest movement of humans ever out of their usual environment, strongly influencing local, regional, national and international economies, and is responsible for about 10% of world's domestic product. However, regardless the crucial role of tourism, there is a lack of quantitative considerations of its flows, although it is essential for understanding the self-organization of human traveling patterns, and global wealth net flows. The UNWTO data presented on this section is the largest database of human travel, representing 763 million tourist arrivals.

An important statistical property to directed networks is the reciprocity, which on the tourism network means the appetency to exchange tourists. The links in the network are composed by 10% bidirectional links and 30% of asymmetric links. If country B as tourist arrivals from country A, then the probability that country A as tourist arrivals from B is only $\frac{1}{4}$, so the network is significantly directed. There is also a large possibility of new connections, since 60% of all the pairs of countries are not connected to one another. It is also concluded that the outbound tourism particularly benefits from market diversity.

It is also shown the first empirical results of the benford to the inbound and outbound tourism on an international scale. This result not only supports the reliability of the secondary data (of the UNWTO) but also proposes that tourist arrivals can be simulated by using the benford law.

7 Scaling Law of Human Traveling

Introduction

The structure and evolution of the world tourism network has been evolving or by a matter of chance or by some kind of rules that naturally emerge on the competitive market place (see models on section 3.3). On this section the evolution of the network is discussed as well as market concentration (section 7.2) and information flows (section 7.3). The network analized was described on section 6.1, and the significance of the results or theory supporting the basis the theoretical background is described on section 3.

7.1 Network: Random or Scale-Free?

A fundamental aspect of real-world networks is the degree, representing the number of connections (links) of each party (node). In random graphs nodes have similar degree, although many real-world networks have some nodes that are significantly more connected than others, many of those are scale free, having connectivity distributions that decay as a power-law. The preferential attachment means that nodes with high degree are preferential.

In general which are the major differences between random networks and scale-free networks (see Fig. 7.1). For random networks, most nodes have approximately the same number of

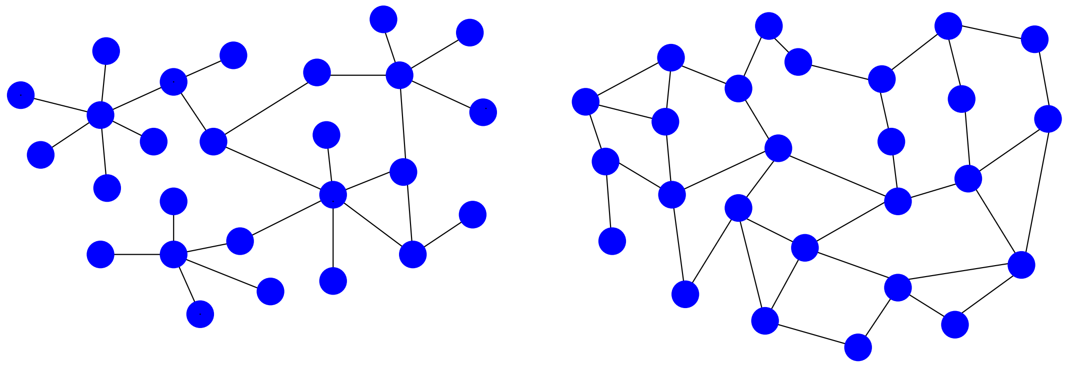


Fig. 7.1: Random network versus scale-free network.

links, $k \approx \langle k \rangle$, where $\langle k \rangle$ represents the average value. The exponential decay of $P(k)$ guarantees the absence of nodes with significantly more than $\langle k \rangle$ links. In contrast, the power-law distribution implies that there is an abundance of nodes with only a few links, and a small - but significant - minority that have a very large number of links.

We can imagine the highways crossing cities in a country. The number of highways crossing a given city is always around a certain number. Bigger and smaller cities may have the same number of crossing roads, although they diverge much on the volume. In contrast, scale-free are more like the world air network [Guimera et al., 2005]. On this network the most airports are served by a small number of carriers, and a few hubs, such as Chicago or Heathrow, from which links emerge to almost all other airports.

On the worldwide tourism network two networks are represented. One considering the topology of the connections. For example, Portugal is connected to Spain if there are tourists (independent of the absolute number) from Portugal to Spain. On the other way around, Spain connects to Portugal if there are tourists going from Spain to Portugal. The overall network represents the existence of connection for every two countries. The second network, called the weighted network, is the one where to each connection is added a weight regarding the number of tourist from one country to another. Let us say that it represents the volume of a given connection.

Network's topology displays the degree distribution $P(k)$, probability that a node has degree k, which applied to world tourism network are studied two degree distribution functions,

$P_{in}(k)$ representing the probability that a node has k nodes directed to itself (probability of countries with tourism from k *inbound* countries), $P_{out}(k)$ representing the probability that a node has a total of k edges to other nodes (probability of countries with tourism to k *outbound* countries). Most networks have a scale-free degree distributions [Albert and Barabási, 1999], which have a power-law tail $P(k) \sim k^{-\theta}$. These measures give us insight into the various roles and groupings in a network – who are the connectors, leaders, bridges, isolates, who is in the core of the network, and who is on the periphery.

Definition 7.1 *The **in-degree** of a country is an indicator of its attractiveness has a destination country, **destination attractiveness indicator**, which increases with the number of origin countries that have flow of tourists to the destination on analyzes.*

Definition 7.2 *The **out-degree** of a country is an indicator of its emanation has a tourism origin country, **destination emanation indicator**, which increases with the number of countries that thecountry on analyzes has flow of tourists to.*

An exponential network is provided by the usual random graph, with $P(k)$ decreasing exponentially fast, although scale-free networks display a hub-like hierarchies, with $P(k)$ decreasing as a power-law [Albert and Barabási, 2002]. In our case, the *in* and *out-degree* distributions decrease exponentially fast, cumulative distribution functions. On figure 7.2 (a) and (b), respectively $P_{in}(k)$ and $P_{out}(k)$. The topological network does not displaying scale-free behavior, similar result on de Montis et al. [2007]. To have a more precise analyzes of the network is performed the weighted analyzes.

The probability distribution function of the weights, $P(w) \sim w^{-\tau}$ has a power–law behavior, with exponent $\tau = 1.55$, see figure 7.3.

On the average the shortest path length between countries is $l = 1.84$, and the diameter is 4, which are small values in accordance with a small–world behavior $l \sim logN$. The shortest path length and diameter are specially small on the worldwide tourism network.

Summary 7.1 ***Small-world:*** *This means that any two countries have a high probability of being*

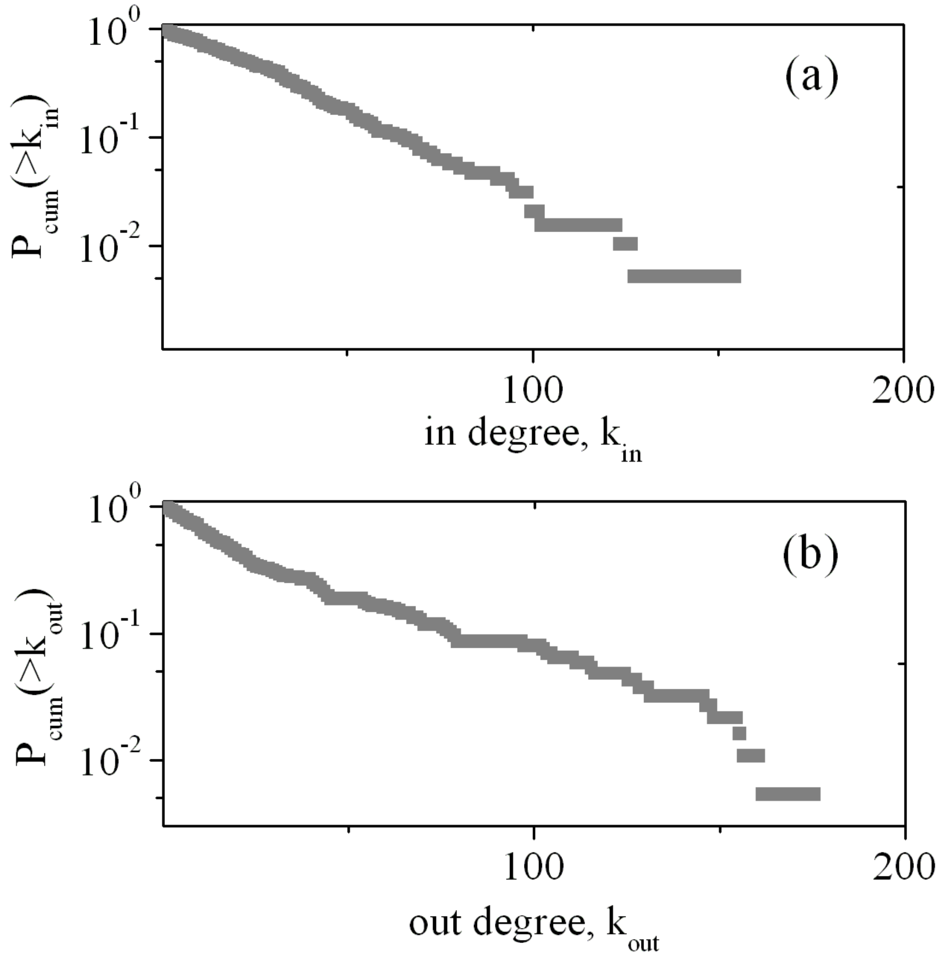


Fig. 7.2: The connectivity of inbound and outbound tourism show a random network behavior: $(a) P_{in}(> k_{in})$, log-linear. (b) $P_{out}(> k_{out})$, log-linear.

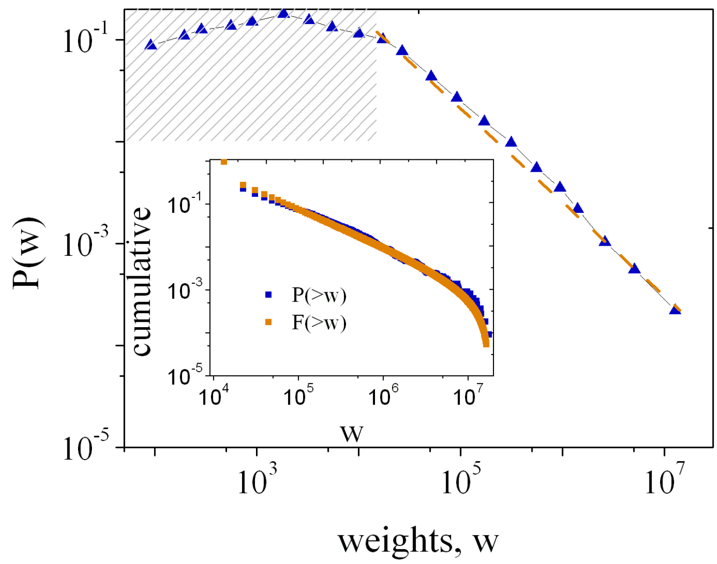


Fig. 7.3: The scale-free behavior of tourism flows: Log-log plot of the weight distribution, $P(w) \sim w^{-\tau}$, where $\tau = 1.55$.

themselves connected, or that have very few intermediate country through each a connection is present. The small-world property has strong influence on the dynamics of the network, like spread of information, innovation, knowledge, promotion, or any other propagation process.

The tourism international network is a giant component, so that all countries have a path or paths to any of the other countries. The fact of being a giant network and having a small shortest path length can imply fast transferring of knowledge and information.

It is also relevant to study the strength of the nodes, which on a directed network each node has strength *in*, $s_{in}(i)$ (eq. 7.1), and strength *out*, $s_{out}(i)$ (eq. 7.2). It measures the strength of the nodes on relation to the total weight of their connections. On the world tourism network, strength *in* represents the *inbound* tourism, and strength *out* represents the *outbound* tourism. Strength is a measure of centrality for weighted networks:

$$s_{in}(i) = \sum_{j \in v(i)} w_{ij}, \tag{7.1}$$

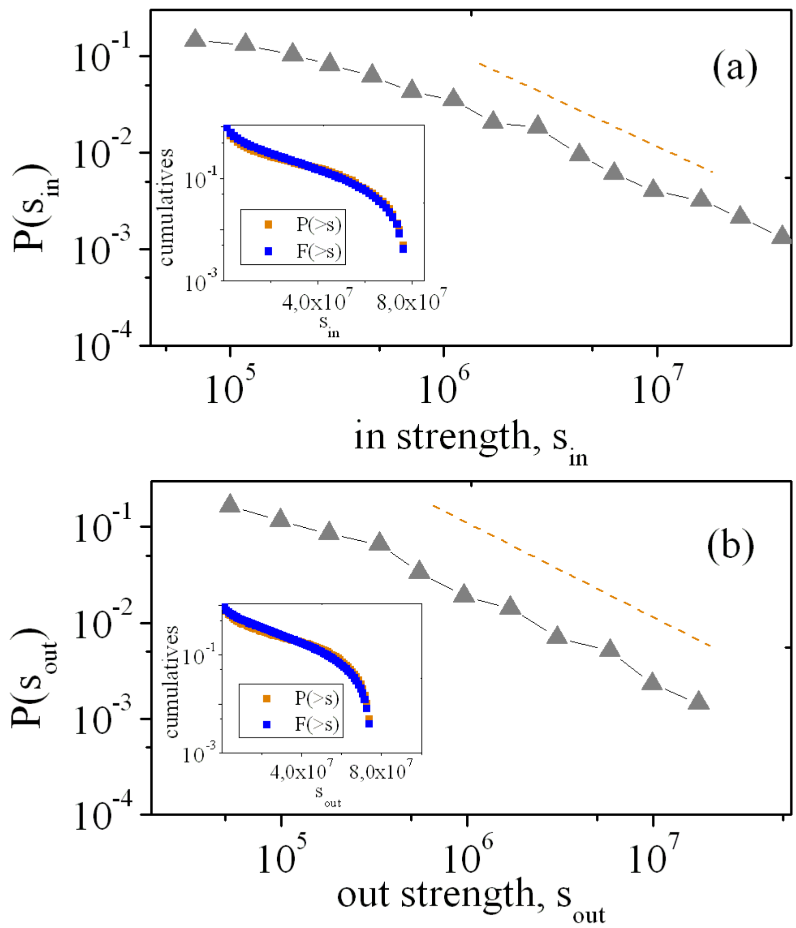


Fig. 7.4: Scale-free behavior of inbound and outbound tourism. Log-log plot of strength *in* distribution, $P(s_{in}) \sim s_{in}^{\theta_{in}}$ where $\theta_{in} = 0.98$ and strength *out* distribution $P(s_{out}) \sim s_{out}^{\theta_{out}}$, where $\theta_{out} = 0.92$.

$$s_{out}(i) = \sum_{j \in \upsilon(i)} w_{ji}. \tag{7.2}$$

The strength *in* distribution and strength *out* distribution functions are also fitted by a power-law, respectively $P(s_{in}) \sim s_{in}^{\theta_{in}}$ and $P(s_{out}) \sim s_{out}^{\theta_{out}}$, where $\theta_{in} = 0.98$ and $\theta_{out} = 0.92$, represented on figure 7.4. The exponents of $P(s_{in})$ and $P(s_{out})$ are smaller than 2. Notice that $P(s_{in})$ has exponent close to 1, and $P(s_{out})$ has exponent close to 0.9.

Summary 7.2 ***Scale-Free:*** *A power-law behavior of $P(w)$, $P(s_{in})$ and $P(s_{out})$ have a strong meaning after the structure of the network, describing the way the weights, and strength centrality,* inbound *and* outbound *tourism, are distributed. The weights and strengths range on a large spectrum of values, and the heavy-tailed distribution implies that nodes have a certain probability of having large strength values, where the average of all intermediate values has no meaning.*

The observations on topological and weighted network reveal different structural results, therefore the relation of topological and weighted flows is studied in more detail, $s(k_{in})$ and $s(k_{out})$. The result is depicted on figure 7.5. The function for *in* strength,

$$s(k_{in}) = k_{in}^{\gamma_{in}}, \tag{7.3}$$

where $\gamma_{in} = 1.1$. For $\gamma = 1$ degree and weight are independent [Barrat et al., 2004a]. So $S(k_{in})$ and k_{in} are close to independent, revealing a very small relation between them. On the other side, $s(k_{out})$,

$$s(k_{out}) = k_{out}^{\gamma_{out}}, \tag{7.4}$$

$\gamma_{out} = 1.75$, revealing a strong relation between strength *out* and degree *out*. This means that *outbound* tourism increases with degree *out*.

Interestingly, when analyzing the diversity of the market and its strength, comes out that inbound and outbound tourism have distinguished outcomes. Even so, both have a power-law behavior, and unavoidable fluctuations. The diversification of outbound markets has a strong and positive increase on total outbound tourism meaning that the flow grows faster than the degree. On the relation between the inbound tourism and its market diversification, the two variable are close to independent.

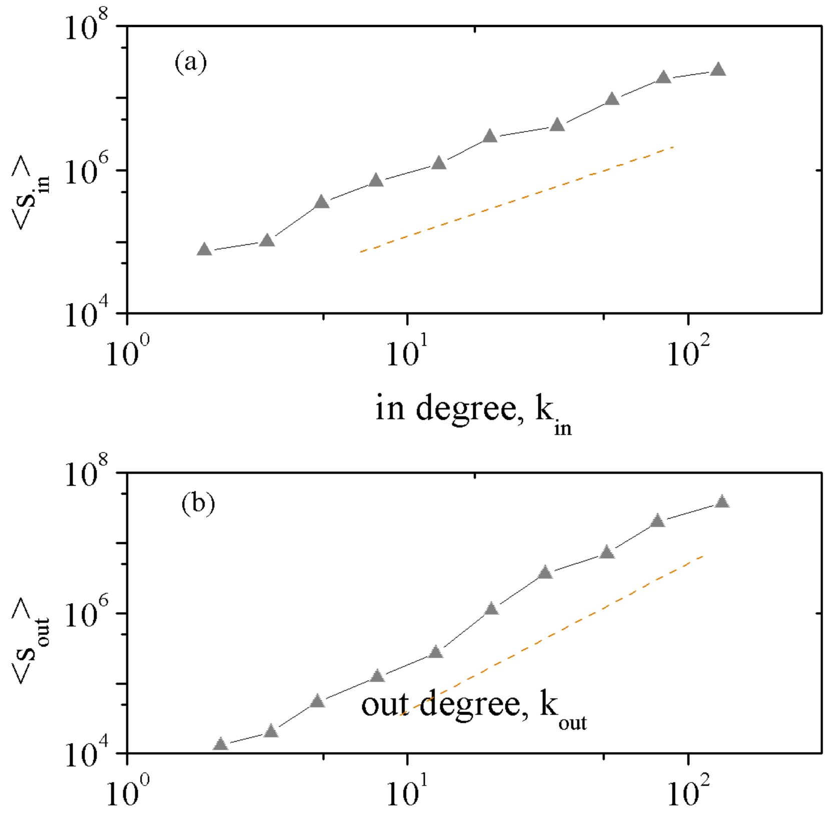


Fig. 7.5: Intensity plays an important role on network behaviour. The relation between degree and strength is closely independent on (a) inbound tourism, $s(k_{in}) = k_{in}^{\beta_{in}}$ with $\beta_{in} = 1.1$, but has a (b) strong relation on outbound tourism, $s(k_{out}) = k_{out}^{\beta_{out}}$ with $\beta_{out} = 1.75$.

7.2 Market Concentration: Disparity Indicator

One important question in tourism is the distribution of the inbound and outbound tourism market. The distribution of the number of tourist from different source markets is an important strategic factor. Some countries have origin markets carrying similar flows, although some others are very diverse in terms of markets dominance. Therefore, it is useful to consider a measure of disparity [Miguéns et al., 2007]:

$$Y(i) = \sum_{j \in \upsilon(i)} \left(\frac{w_{ij}}{s_i} \right)^2 . \tag{7.5}$$

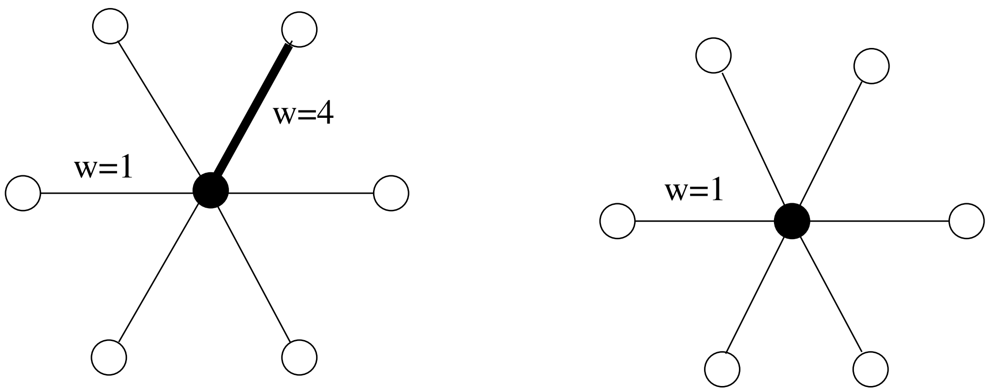


Fig. 7.6: Disparity for the topological and weighted network.

In the same way has for the other measures we average in order of the degree to obtain the $Y(k)$. The results are shown in fig. 7.7, which is a log-log plot. The outbound weights disparity displays a fit approximately $k_{out} * Y(k_{out}) = k_{out}^{\theta}$ and the inbound tourism $k_{in} * Y(k_{in}) = k_{in}^{\theta}$. These results mean that the outbound and inbound markets are very heterogeneous and have few dominant edges with small weights on the other edges.

7.3 Information Flow

So far we studied the local interactions of a node and its neighbors. But which are the prominent nodes on the global pictures? And how important are they when concerning information flow?

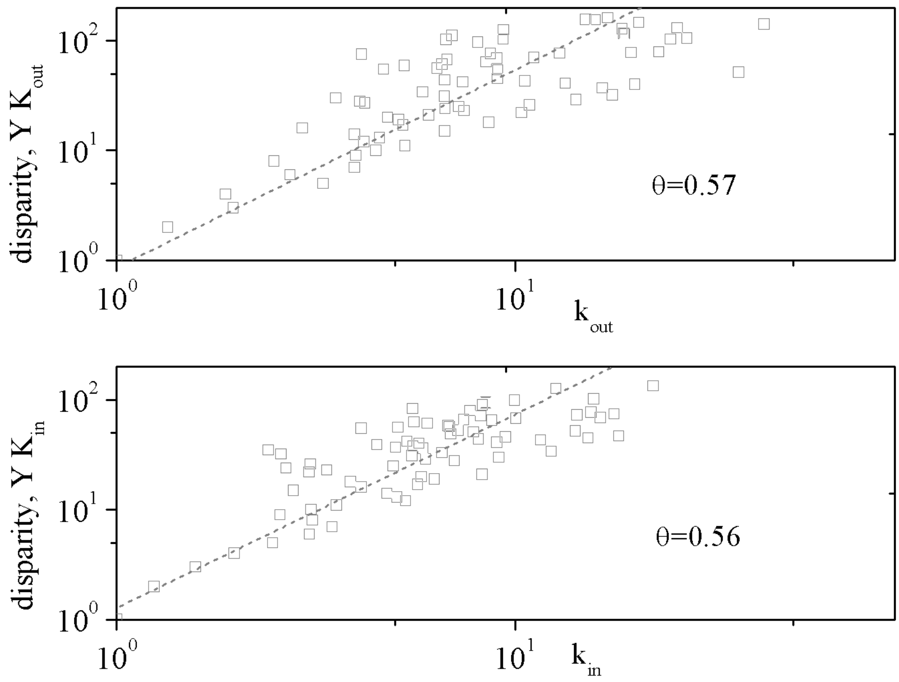


Fig. 7.7: Is the inbound and outbound market dominated by a few connections? It is displayed a log-log plot of out disparity versus out-degree (on the top) and in disparity versus in-degree (down plot). Out disparity $k_{out}Y(k_{out}) = k_{out}^{\theta_o}$ for $\theta_o = 0,57$ and in disparity $k_{in}Y(k_{in}) = k_{in}^{\theta_i}$ for $\theta_i = 0,56$.

Common wisdom in personal networks is ”the more connections, the better,” although this is not always true. Some connections can be more crucial than others bridging to otherwise disconnected nodes. Like represented on Fig. 3.5 (section 3.2.3) the nodes carrying more infromation are not necessarily the ones with more connections.

To measure the traffic crossing each of the countries was performed the Dijkstra’s algorithm, as described on section 3.2.3. The cumulative distribution of both unweighted and weighted network are ploted on Fig. 7.8. The straight lines of the distribution function means that betweeness also follows a power-law $P(b) = b^{\theta}$.

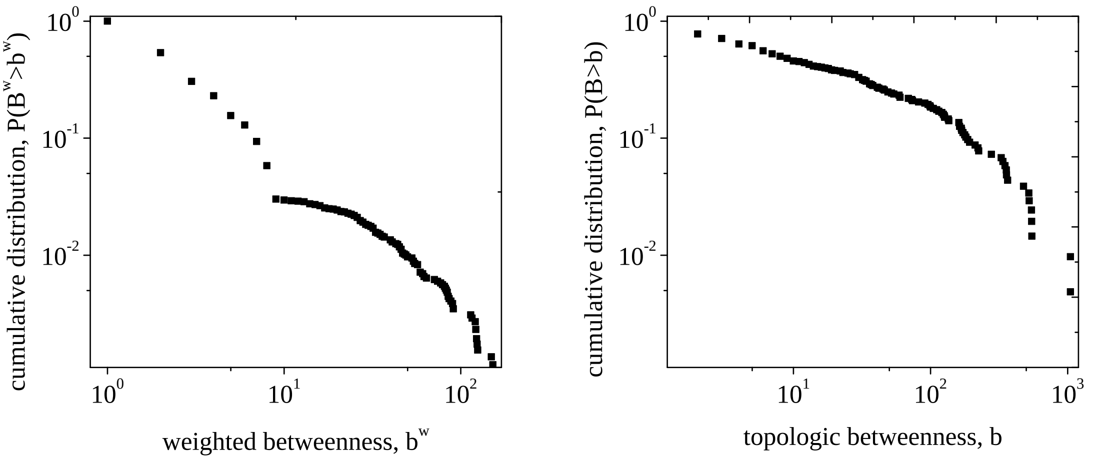


Fig. 7.8: The betweeness centrality is the most used measurement of information flow.

7.4 Conclusion

Two networks are analyzed, one representing the overall connections (degree) of tourism connections and another representing the absolute number of arrivals (weight) on a international scale. These two nets are considerable different and have different implication.

The tourism network, considering the degree of different countries, is similar to a random network [Miguéns and Mendes, 2008b]. This is in contrast with the worldwide air network, observed to be a scale-free [Guimera et al., 2005]. On the worldwide tourism, degree distribution measures the probability of the number of incoming connections $P(k_{in})$, for the inbound tourism, and the probability of the number of outgoing connections $P(k_{out})$, for outbound tourism. The cumulative distributions on both degrees are depicted by a log-normal fitting, so the degree distributions display an exponential decay, typical of a random network [Erdös and Rényi, 1960]. The degree distributions decay is greatly faster than the power-law degree distribution depicted in other social, technological [Loudhouse-Research, 2007] and economic [Serrano et al., 2007] networks [Caldarelli et al., 2004; Fagiolo, 2006; Dorogovtsev and Mendes, 2003; de Montis et al., 2007].

A scale-free network is obtained when we have in consideration the travelers traffic (weight) between countries. Accordingly, the inbound and outbound distributions reasonably suggest a growth model in which new connections are chosen as a result of preferential attachment [Miguéns and Mendes, 2008b], where higher degree countries are more likely to attract new connections. While, plausibly, destinations are chosen randomly, not being clear if there is any advantage for making new connections with high in-degree countries.

The findings reveal in this way that the tourism network has an homogeneous topology, meaning that countries have approximately the same number of connections. Even if the arrivals of tourists can be dramatically different, the countries have similar number of international tourism origins.

In contrast, the network present shows a scale-free behavior on the weights, inbound and outbound tourism. The number of tourists on a destination increase the probability of having new incomer tourists. Also the other way around, the number of tourists on a origin country that travel abroad increase the probability of having new outgoing tourists. The popular countries are the ones that more easily add new connections. And a country with more tourists traveling abroad more easily has more travelers.

The random versus scale-free network means that, the network changes on the preferential attachment of new tourists, but does not change on the topological relation. This notion of preferential attachment can be some how related with a sort of a social phenomena, where travelers or tourists take into consideration word-by-mouth, marketing, travel promotion, when choosing a destination. This works in both ways, for inbound and outbound tourism. The preferential attachment seems to work on the direction of ”fame” of some countries, but not necessarily on the number of connections.

It is remarkable how, on the tourism network structure, the consideration of weight is crucial, converting a random network of connectivity into a scale-free network of flows. So far, recently reported technological and social networks have a degree distribution with a scaling exponent ranging from 2 to 3. However, exponential degree distributions were also reported on an email network Ebel [2002] and on a structure of inter-urban traffic [de Montis et al., 2007].

So, two consequences are expected, the network is growing due to a scaling up, with an increase of flows intensity and/or due to a scaling out by new connections between countries. It is worth noticing that the cut-off limit on the distribution is explained by the countries capacity constraints that limit the ability of the network to scale up. The factors limiting tourism supply tend to be political, lack of security, etc. Similarly the distribution of inbound tourism and outbound tourism also display a power-law.

On the average the shortest path length between countries is $l = 1,84$, and the diameter is 4, which are small values in accordance with a smallworld behavior $l \approx logN$. The shortest path length and diameter are specially small on the world tourism network. This means that any two countries have a high probability of being themselves connected, or that have very few intermediate country through each a connection is present. The small-world property has strong influence on the dynamics of the network, like spread of information, innovation, knowledge, promotion, or any other propagation process. The tourism international network is a giant component, so that all countries have a path or paths to any of the other countries. The fact of being a giant network and having a small shortest path length can imply fast transferring of knowledge and information, or destination image propagation.

A question arise when we think about the proliferation of scale-free networks as a model of travel and tourism industry, and the increasing dependence on them (particularly given their prevalence in energy, transportation, and communications systems): how reliable are these networks?

Scale-free networks are tolerant of random failures? In a random network, a small number of random failures can collapse the network. A scale-free network can absorb random failures up to 80$\%$ of its nodes before it collapses. The reason for this is the inhomogeneity of the nodes on the network – failures are much more likely to occur on relatively small nodes. Scale-free networks are extremely vulnerable to intentional attacks on their hubs. Travel and tourism industry is directly affected by other factors, like terrorism, war and natural disasters. These factors are randomly affecting different parts of the world. It this way, the travel and tourism industry is rapidly able to recover from those ”attacks”.

A possible intentional attack to a country, as a tourist destination, can have strong negative effects. Scale-free networks are extremely vulnerable to intentional attacks on their hubs: Attacks that simultaneously eliminate as few as 5-15% of a scale-free network's hubs can collapse the network. Simultaneity of an attack on hubs is important. Scale-free networks can heal themselves rapidly if an insufficient number of hubs necessary for a systemic collapse are removed.

Travel and Tourism, an Economic Network? 8

Introduction

On this section is considered the way countries couple with one another. How is a tourist destination coupled with its similar? It is presented some indicators of nonsocial behavior that show how tourist destinations are related. Are destinations connected with each other on a preference way? Is there a mechanism that makes some connections more probable than others. In a social system is usually observed assortative mixing behavior, observed when the nearest neighbors of nodes with highly connected are also highly connected. On economic, technological and biological systems are generally observed disassortative mixing, observed when the nearest neighbors of nodes with high degree have low degree.

The degree-degree correlation (see section 3.2.5) and clustering coefficients (see section 3.2.4) are analyzed on the unweighted and weighed world tourism network. The assosrtativity is depicting the way countries couple together and the clustering reveals possible mechanisms of social, biological, economic or other nature that acts has an organizational principle.

To probe the weighted networks' architecture was introduced the clustering coefficient for weighted networks. It is a conceptual challenge because of the highly heterogeneity of weights (flows) and also for being a directed network. Capturing the patterns of clustering in a weighted network is a new concept, and different methods have been proposed [Barrat et al., 2004a; Park and

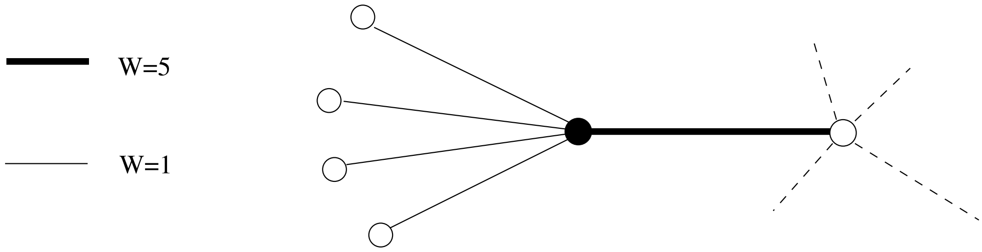


Fig. 8.1: Assortativity: the unweighted and weigthed degree-degree correlations.

Kim, 2006; Onnela et al., 2005; Zhang and Horvath, 2005; Holme et al., 2007]. A critical review of these methods have been presented [Saramaki et al., 2007]. To better understand the clustering pattern on the network more than one proposed definition is studied, which give rise to different aspects of the clustering property.

8.1 Is There Assortativity on Travelers Choices?

To measure the correlation on the network over degree, one may also study the average nearest-neighbors degree. This measure the tendency of node i to be connected to nodes with the same degree,

$$k'_{nn}(i) = \frac{1}{k_i} \sum_{j \in \upsilon(i)} k_j, \tag{8.1}$$

where $\upsilon(i)$ denotes the set of neighbors of i. Considering that our network is directed, we correlate the *in* degree of node i with the *out* degree of its neighbors:

$$k_{nn}(i) = \frac{1}{k_i^{in}} \sum_{j \in \upsilon(i)} k_j^{out}. \tag{8.2}$$

We can also average the over nodes of the same degree:

$$k_{nn}(k) = \frac{1}{NP(k)} \sum_{k_i = k} k_{nn}(i). \tag{8.3}$$

This measure is also called *associative mixing* if nodes with high degrees have most of their neighbors with high degrees, represented by a growth of $k_{nn}(k)$ with k. For a decreasing of $k_{nn}(k)$ with k it is denominated disassortative mixing. This happens when nodes with high degrees have mainly neighbors with low degree. The world tourism network displays disassortative mixed. This behavior is mostly detected on transportation networks, providing a pattern where the hubs connect to the small degree nodes at the periphery of the network [Newman, 2003a].

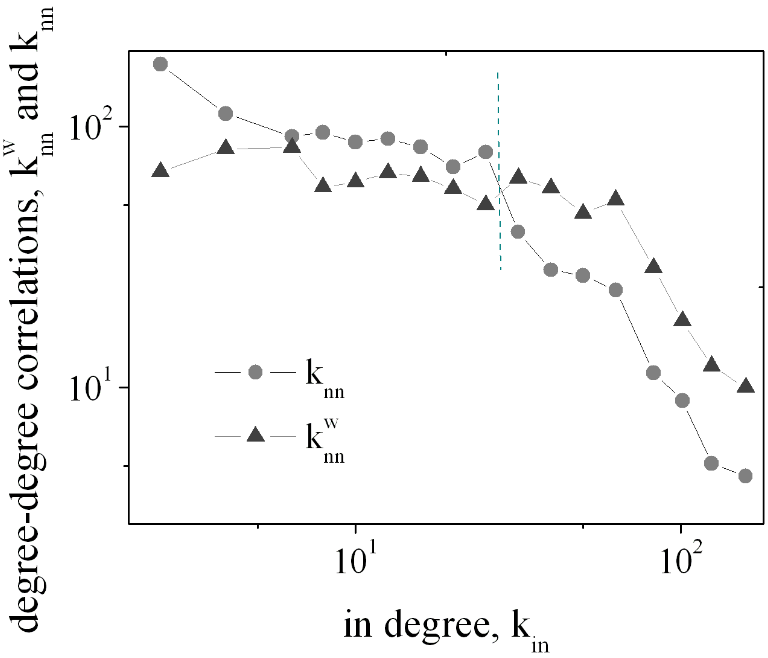


Fig. 8.2: The tourist destinations have an assortative behavior, on the connectivity and weighted network. It is displayed the log-log plot of *in* degree - *out* degree correlations, unweighted $k^w_{nn}(k)$ and weighted $k_{nn}(k)$, versus k_{in}. For low degrees $k^w_{nn}(k) < k_{nn}(k)$ and for high degrees $k^w_{nn}(k) > k_{nn}(k)$.

Degree-degree correlation for a weighted network is given by de Montis et al. [2007]:

$$k^w_{nn}(i) = \frac{1}{s^{in}_i} \sum_{j \in \upsilon(i)} w_{ji} k^{out}_j, \tag{8.4}$$

$k^w_{nn}(k)$ measures the local weighted average of neighbors degree, see Fig. 8.1. The spectrum of world tourism network on topological (equation 8.3) and weighted degree-degree correlations (equation 8.4) if represented on figure 8.2. For $k^w_{nn}(k) > k_{nn}(k)$ the edges with the larger weight are directed to the neighbors with larger degrees, and $k^w_{nn}(k) < k_{nn}(k)$ the edges with the larger

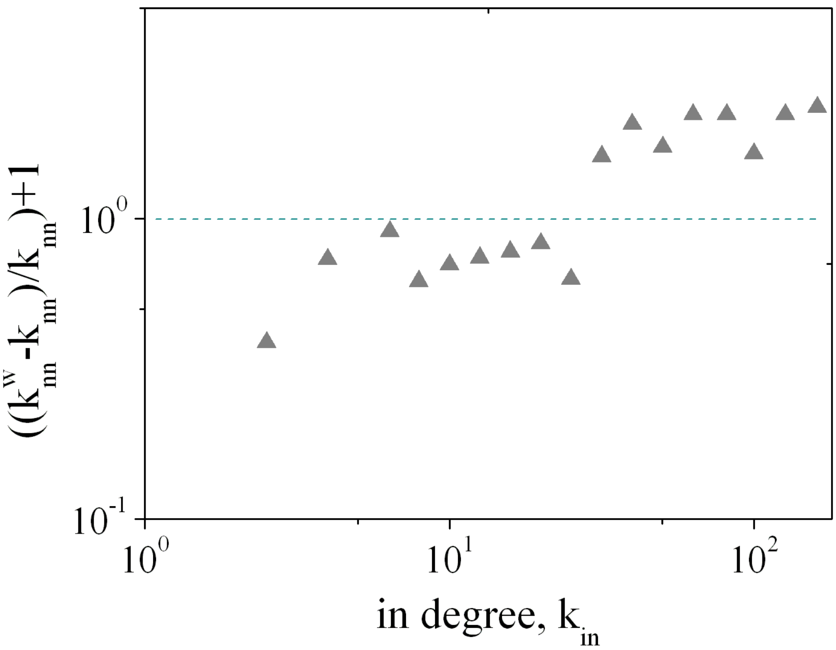


Fig. 8.3: Comparing weighted and topological degree correlation. For low degree $k^w_{nn}(k) < k_{nn}(k)$ and for high degree $k^w_{nn}(k) > k_{nn}(k)$, meaning that low (high) degree nodes have their edges with large weight directed from nodes with low (high) degree.

weight are directed to the neighbors with lower degrees [de Montis et al., 2007], see Fig. 8.2. The weighted degree-degree correlation is slightly decreasing (see Fig. 8.3), following the same behavior as the topological correlation, but with a slower slop. For low degrees $k_{nn}^w(k) < k_{nn}(k)$ and for high degrees $k_{nn}^w(k) > k_{nn}(k)$, meaning that low degree nodes have their edges with large weight directed from nodes with low degree, and high degree nodes have their edges with large weight directed from nodes with high degree.

8.2 Nonsocial Travelers Behavior and Local Triangulations

The clustering coefficient, for the tourism network, as a first attempt to measure cohesiveness is defined as the probability of two countries that are directly connected to a third country are also directly connected to each other. Network theorists first studied topology for undirected and unweighted networks, defining clustering coefficient [Watts and Strogatz, 1998; Barrat and Weigt, 2000] of a node i as:

$$c(i) = \frac{2E}{k_i(k_i-1)}, \tag{8.5}$$

where i is the node, k_i is the degree of node i, and E is the number of edges between neighbors of node i. $C(i)$ belongs to the interval $[0, 1]$ and gives the local connection of the network. $C_i = 0$ if any of the neighbors of node i are connected, and $C_i = 1$ if all the neighbors are connected. Averaging the nodes over the network with the same degree we obtain $C(k)$ (on Fig. 8.5). The global clustering is given as $C = \frac{\sum_{i=1}^{N} C(i)}{N}$, where $C = 0.655$.

The first introduced weighted clustering coefficient was [Barrat et al., 2004a]:

$$C^w(i) = \frac{1}{s_i(k_i-1)} \sum_{j,h} \frac{w_{ji} + w_{hi}}{2} a_{ji} a_{hi} a_{jh}, \tag{8.6}$$

where a_{ij} a_{ik} and a_{jk} belong to the adjacency matrix. $s_i(k_i - 1)$ is the normalization factor and ensures that $C^w(i)$ belongs to the interval $[0, 1]$. Averaging for the nodes with the same degree we obtain $C^w(k)$.

The weighted clustering measures the cohesiveness of local triplets, considering the

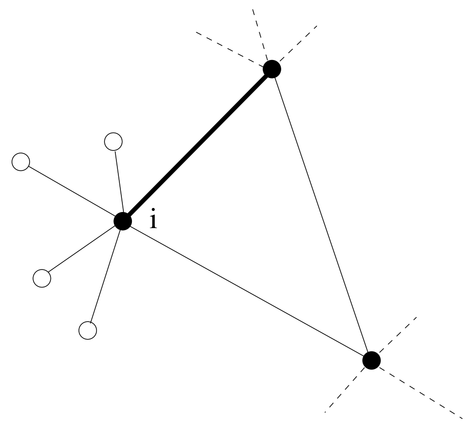


Fig. 8.4: The larger countries are well interconneted. The weighted and topologic clustering relation.

weight of the connection to node i. This way we are accounting for the relative weight of the triplets to the strength of the node. Notice that $C(i) = C^w(i)$ when weights are binary.

$C(i)$ and $C^w(i)$ have both a continuously decaying spectrum Figure 8.5, showing that hubs have a much lower clustered neighborhood than low degree nodes. The countries with low number of travel destination belong to well interconnected communities, among which exchange travelers. while countries with a large number of travel destinations function has hubs connecting other regions.

Countries with high number of travel destinations are stronger on edges with higher weights ($C^w(k)/C(k) \approx 1.2$), meaning that there is a tendency to agglomerate flow on interconnected groups. Notice that for low degree $k < \langle k \rangle$, ($C^w(k)/C(k) \approx 1.05$), and then $C^w(k)/C(k)$ grows with degree, the higher is the degree the more tendency to have stronger connections with other high connected countries.

The spectrum variation of C^w is much smaller than on C. Over all the spectrum countries tend to interconnected group with high weight links. Therefore, countries with high degree nodes tend to have travelers with other high degree countries, also called the rich club phenomenon [Zhou and Mondragón, 2004].

The weighted clustering C^w, includes two of the three links of the triangle. To measure the

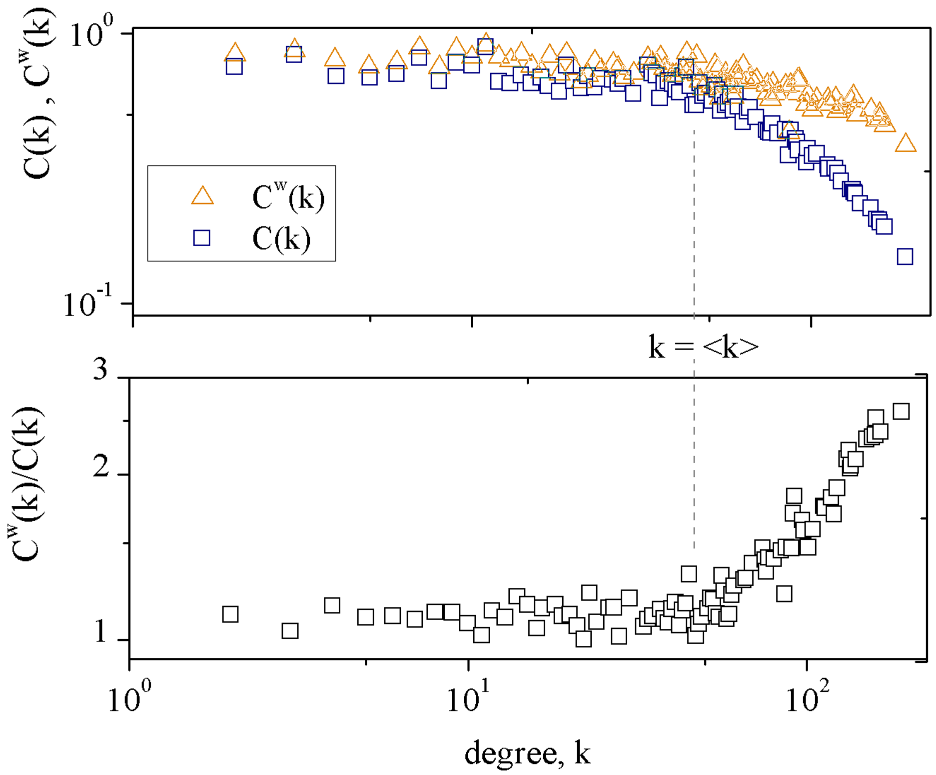


Fig. 8.5: The clustering coefficient over degree, $C(k)$ and $C^w(k)$. Notice that $C^w(k) > C(k)$, meaning that highly connected countries tend to exchange more travelers within interconnected groups. For countries more connected they the average $k > \langle k \rangle$, $C^w > C$ is more evident, high degree countries tend to form interconnected groups with high-weighted connections.

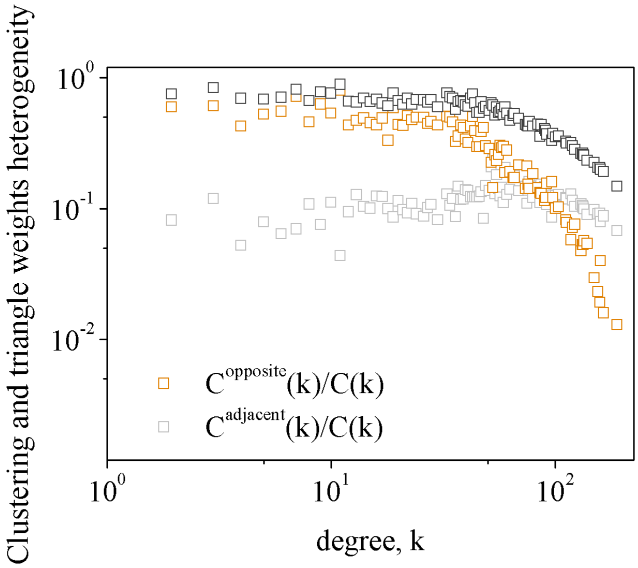


Fig. 8.6: Comparation of the clustering coefficients: C^w, C^{edge} and $C(k)$. Source: [Miguéns and Mendes, 2008a]

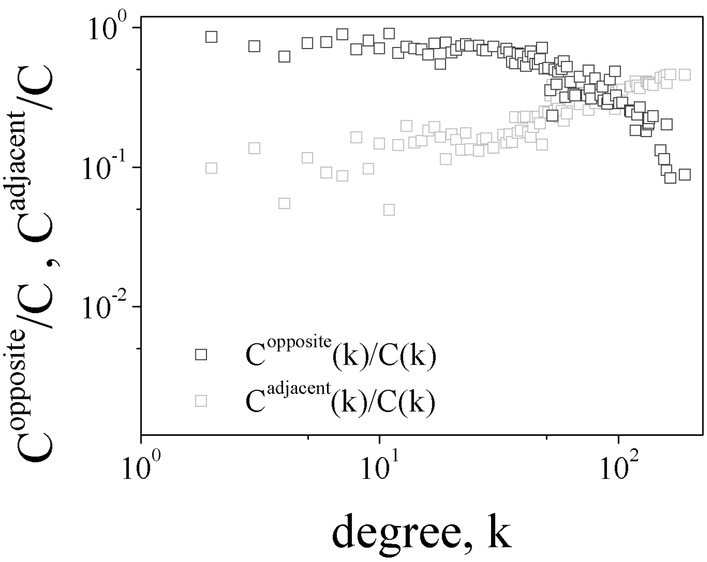


Fig. 8.7: On relation with the clustering both other clusterings, edge and opposite, increase faster. Source: [Miguéns and Mendes, 2008a].

contribution of the third link eq. 8.7, are compared the weighted of the third link in the maximum of the other two:

$$C^{edge}(i) = \frac{1}{k_i(k_i-1)} \sum_{j,h} \frac{w_{jh}}{Max(w_{ij}, w_{hi})} a_{ji} a_{hi} a_{jh}. \tag{8.7}$$

$C^{edge}(i)$ belongs to the interval $[0,1]$, is $C^{edge}(i) = 0$ if any of the neighbors of node i are connected, and $C^{edge} = 1$ if all the neighbors are connected. Averaging over nodes with the same degree we obtain $C^{edge}(k)$. For a unweighted network $C^{edge}(k) = C(k)$.

8.3 Yearly Competitiveness and Global Economic/ War/ Natural/ Terrorist Factors

The relation between topological and weighted clustering, for undirected networks, based on equation 8.6 has been studied in some real world networks, like urban movements, scientific collaboration network and worldwide airport network [de Montis et al., 2007; Barrat et al., 2004a].

Clustering on studied real world networks decrease on degree k, suggesting that low degree nodes belong to well connected communities, and high degree nodes connect different communities. For example, this has been studied in urban movements [Chowell et al., 2003; de Montis et al., 2007].

The worldwide tourist arrivals/departures network is a random network on its topology and a scale-free network when considering the weighted network [Miguéns and Mendes, 2008b]. So, we would expect the clustering coefficient for the unweighted network to follow the clustering for random networks [Ebel, 2002; Newman and Park, 2003], given by

$$C_{RG} = \frac{(\langle k^2 \rangle - \langle k \rangle)^2}{N \langle k \rangle^3}. \tag{8.8}$$

In accordance with the expected $C_{RG} = 0.65$ (eq. 8.8) we get the value of $C = 0.655$ for the world tourism network. This means that a generalized random network is a suited model of the clustering behavior of the worldwide tourism flows. Newman and Park [2003] showed a model of for small networks and large nonsocial networks, like the internet [Maslov et al., 2004] and a food web of organisms [Martinez, 1991] having a clustering approximately by the one for the random network model.

By other hand, empirically calculated clustering for real-world social networks are higher than the generated by the random network model, like a film actor collaborators network [Amaral et al., 2000] and an email network [Newman et al., 2002]. Therefore, the world tourism network has a clustering pattern of nonsocial networks, the same pattern expected on a random chance of linkages. In accordance with previous work, the trajectories of travel flows do agree with nonsocial networks behavior, remarking the nonsocial patterns of travelers [Miguéns and Mendes, 2008b].

Summary 8.1 *On the world tourist arrivals network the clustering coefficient was found to depict the evolution dynamics affected by global economic factors. The world tourism network flows have a yearly growth rate of* 3.7%, 1999 *a* 2004*. However, economic recession, geopolitic factors, diseases and terrorism attacks have slowed down the flows on the beginning of this century,* 2001 *and* 2003*. In* 2001 *tourism slowed down due to September* $11'$*s, and in* 2003 *was the biggest slowed down in tourism, affected by Iraq war, SARS and the prevailing weak economy.*

On 2004 tourist arrivals recovered and reached a very high grows of $10,7\%$ over 2003. Clustering coefficient depicts the world tourism flow dynamics, and the economic factors. A similar study over crash on a financial network was performed by Onnela et al. [2005]. On the world tourism network $C_O^{WDN}(k)$ is calculated for the years of $1999-2004$, see the following Tab. 8.1. The growths evolution of the world arrivals is depicted by the clustering coefficient. The clustering C^w takes into account the two adjacent links of a given node i, and their weights. Although the triangles, let say i, j and k (see Fig. 8.9), can be strongly or weakly connected depending on the homogeneity of the link weights. The neighbors of a node can have a stronger or weaker connection than the node with the neighbors. To capture the relation between the adjacent and opposite links are introduced equations 8.9 and 8.10:

$$C^{opposite}(i) = \frac{1}{k_i(k_i-1)} \sum_{j,h} \frac{w_{jh}}{Max(w_{ji}, w_{hi}, w_{jh})} a_{ji} a_{hi} a_{jh}. \tag{8.9}$$

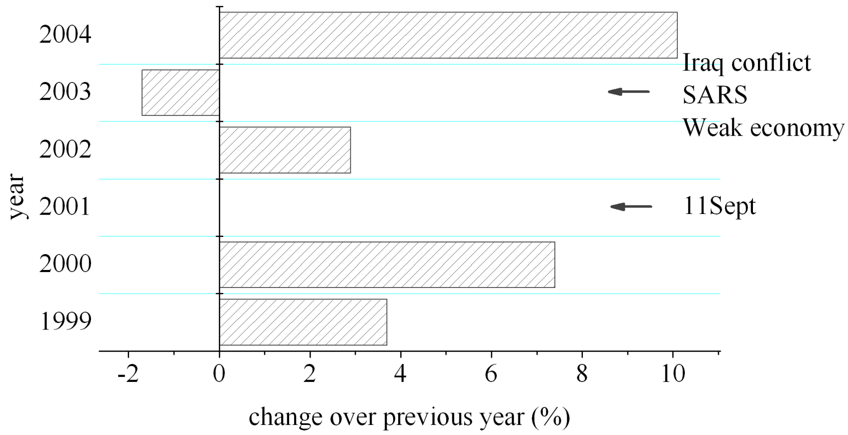
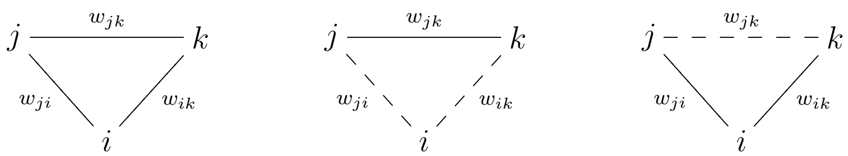


Fig. 8.8: International tourism arrivals on the years 1999 to 2004.



(a) Similar weights, $w_{ji} \approx w_{jk} \approx w_{ik}$.

(b) Higher weight on opposite link.

(c) Higher weight on adjacent links.

Fig. 8.9: The triangle j, k and i can have high heterogeneity of weights, influencing how strongly the triangle is connected.

$$C^{adjacent}(i) = \frac{1}{k_i(k_i-1)} \sum_{j,h} \frac{\frac{w_{ji}+w_{hi}}{2}}{Max(w_{ji}, w_{hi}, w_{jh})} a_{ji}a_{hi}a_{jh}. \tag{8.10}$$

$C^{opposite}(i)$ and $C^{adjacent}(i)$ belong to the interval $[0, 1]$, and equal zero if any of the neighbors of node i are connected. Averaging the nodes over the network with the same degree, k, we obtain respectively $C^{opposite}(k)$ and $C^{adjacent}(k)$. For a unweighted network one gets for both $C^{opposite}(k) = C(k)$ and $C^{adjacent}(k) = C(k)$.

A country with two of its neighbors, when forming a triangle, tend to have a high heterogeneity of weights. It can be that a country is highly connected to its neighbors, but its neighbors have a low connectivity between them, like in Fig. 8.9 (c). Or the opposite, the neighbors can be more connected than the country with its neighbors Fig. 8.9 (b). An homogeneity of weights would happen with all the weights are of the same order, like in Fig. 8.9 (a).

On the acquaintances triangles the hubs have very strong connections with their neighbors (Fig. 8.7), which means that prominence countries commute significantly travelers with their neighbors. Although their neighbor are weakly connected, revealing the importance of hubs on connecting different destination regions. By other hand, countries with low number of connections have highly connected neighbors, higher than the connection with the neighbors.

Onnela et al. [2005] proposed another definition of clustering coefficient:

$$C^w_O(i) = \frac{1}{k_i(k_i-1)} \sum_{j,h} (\hat{w}_{ij}\hat{w}_{ih}\hat{w}_{jh})^{\frac{1}{3}}, \tag{8.11}$$

$\hat{w}_{ij} = \frac{w_{ij}}{max(w)}$, where $max(w)$ is used for normalization. Triangles with edge weight equal to $max(w)$ contribute unity to the sum, and triangles with small weight will have a small contribution. $C^w_O(i) \in [0, 1]$ and reflects how large triangle weights are compared to network maximum. The clustering coefficient over all nodes is $C^w_O = \frac{\sum_{i=1}^N C^w_O(i)}{N}$.

Tab. 8.1: Worldwide tourist arrivals and the clustering on the years $1999 - 2004$.

year	arrivals	grow	grow	$C_d^w(\times 10^{-3})$	C (≈)
1999	639.6				
2000	687.0	↗		$2,374$	$0,654$
2001	686.7	↘	↗	$2,437$	$0,670$
2002	707.0	↗	↘	$2,393$	$0,659$
2003	694.6	↘	↗	$2,514$	$0,670$
2004	765.1	↗	↘	$2,383$	$0,655$

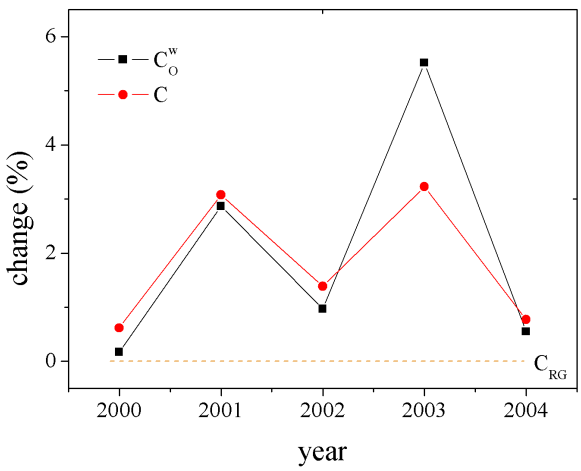
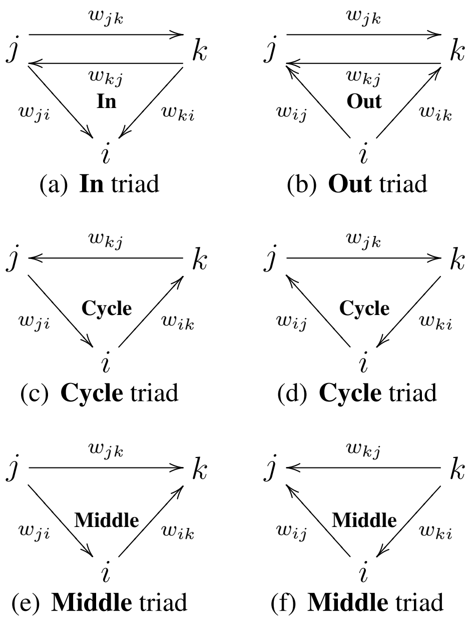


Fig. 8.10: Change (%) on cluster coefficient on the years 2000 to 2004.



(a) **In** triad (b) **Out** triad (c) **Cycle** triad (d) **Cycle** triad (e) **Middle** triad (f) **Middle** triad

Fig. 8.11: Patterns of directedness on clustering: (a) **in** triads: $a_{ji}a_{ki}a_{kj} = 1$ or $a_{ji}a_{ki}a_{jk} = 1$, (b) **out** triads: $a_{ij}a_{ik}a_{kj} = 1$ or $a_{ij}a_{ik}a_{jk} = 1$, (c)-(d) **cycle** triads: $a_{ji}a_{ik}a_{kj} = 1$ or $a_{ki}a_{ij}a_{jk} = 1$, and (e)-(f) **middle** triads: $a_{ji}a_{ik}a_{jk} = 1$ or $a_{ki}a_{ij}a_{kj} = 1$.

8.4 Triangles Diversity, is There a Meaning?

The clustering, C (equation 8.6), refers to a weighted and undirected network. Considering that the tourist arrivals networks is highly heterogenic, the direction of edges should be taken into account. On a directed network there are different patterns of triads around a node, depending on the direction of the edges. On figure 8.11 are depicted the possible patterns of triads, named [Fagiolo, 2006]: *in*, *out*, *cycle*, and *middle*.

The matrix representation of the network allow us to measure different directions of edges by using matrix transpose [Onnela et al., 2005; Fagiolo, 2006]. Generalization of equation 8.11 to directed networks [Fagiolo, 2006] is depicted on the following equations.

The directed and weighted clustering coefficient by Fagiolo [2006]:

$$C_O^{WDN}(i) = \frac{[\hat{W} + \hat{W}^T]^3_{ii}}{2[k^{total}(k^{total} - 1) - 2k^{\leftrightarrow}]}, \tag{8.12}$$

where $\hat{W} = W^{[\frac{1}{3}]} = \{w_{ij}^{\frac{1}{3}}\}$. This clustering coefficient takes into account the weights of all edges in a triangle. On equation 8.6 only the edges from or to the central node are measured.

The patterns are represented on figure 8.11, where triads are divided into four different categories [Fagiolo, 2006]: *cycle* (equation 8.15), *in* (equation 8.13), *out* (equation 8.14) and *middle* (equation 8.16). The pattern *in* (see figure 8.11(a)) is given by:

$$C_D^{in}(i) = \frac{(A^T A^2)_{ii}}{k_{in}(k_{in}-1)}, \;\; \hat{C}^{in}(i) = \frac{(\hat{W}^T \hat{W}^2)_{ii}}{k_{in}(k_{in}-1)}. \tag{8.13}$$

The pattern *out* (see figure 8.11(b) is given by:

$$C_D^{out}(i) = \frac{(A^T A^2)_{ii}}{k_{out}(k_{out}-1)}, \;\; \hat{C}^{out}(i) = \frac{(\hat{W}^T \hat{W}^2)_{ii}}{k_{out}(k_{out}-1)}. \tag{8.14}$$

The pattern *cycle* (see figure 8.11(c) and 8.11(d)) is given by:

$$C_D^{cyc}(i) = \frac{(A)_{ii}^3}{k_{in}k_{out}-d^{\leftrightarrow}}, \;\; \hat{C}^{cyc}(i) = \frac{(\hat{W})_{ii}^3}{k_{in}k_{out}-d^{\leftrightarrow}}. \tag{8.15}$$

The pattern *middle* (see figure 8.11(e) and 8.11(f)) is given by:

$$C_D^{mid}(i) = \frac{(AA^T A)_{ii}}{k_{in}k_{out}-d^{\leftrightarrow}}, \hat{C}^{mid}(i) = \frac{(\hat{W}\hat{W}^T\hat{W})_{ii}}{k_{in}k_{out}-d^{\leftrightarrow}}. \tag{8.16}$$

On Tab. 8.2 are compared the global clustering of the triad patterns. *Cycle* and *middle* do have a prominence in terms of directed and binary clustering, but the weighted analyze reflects the importance of *in*, *out* and *cycle* patterns over *middle*. The difference of values relation over the patterns on the topological and weighted analyze shows the importance of weighted measurements. However an analyze over degree and strength reveal that $\hat{C}^{out}$ is strongly correlated with degree out, k_{out}, and strength out, s_{out}; and $\hat{C}^{in}$ shows no correlation with degree *in* and strength *in*.

8.5 The Coreness of Traveling Flows

The concept of k-core was introduced on section 3.2.6, it has been applied has a strong visualization tool and also as a method to find the nodes that most participate on the functioning of the systems.

Tab. 8.2: Comparing clustering for the different triads patterns: *cycle*, *middle*, *in* and *out*.

Pattern	C_D	$\hat{C}$
cycle	0.214	0.014
middle	0.225	0.001
in	0.132	0.017
out	0.141	0.011

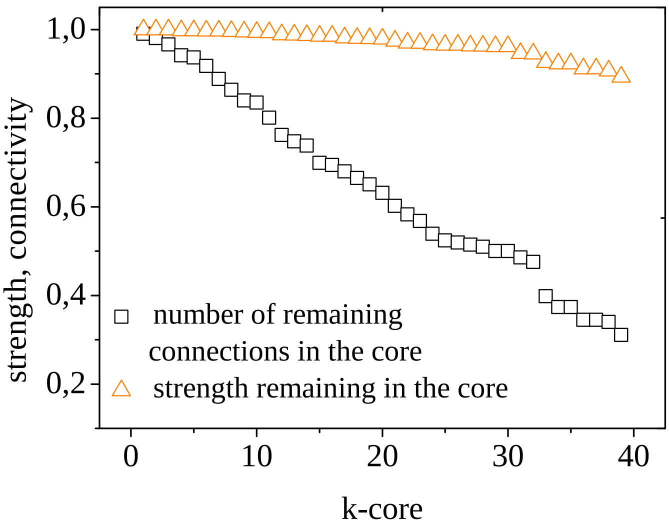


Fig. 8.12: Nodes remaining in subgraph, analyzes on k-core.

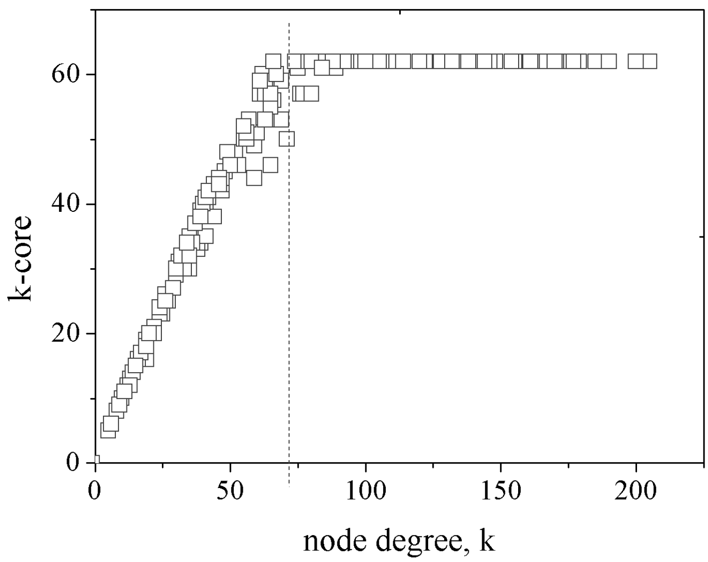


Fig. 8.13: Connectivity in relation with k-core.

In general a k-core is a subgraph in which every node is a neighbor to at least k nodes, representing groups of a graph in which interesting nodes will be found [Seidman, 1983].

On the world tourist arrivals network the k-core decomposition was performed. Remarkably the inner k-core has almost the whole shortest paths, and so most of all the information goes through the inner k-core, resulting on the control of the countries positioned in the inner k-core.

On the decomposition the nodes are removed depending if they maintain a sufficient number of connections inside the subgraph. A question arises on whether most of the connections remain in the subgraphs. By calculting the fraction of remaining connections in the k-cores, as well as the remaining weights, is observed a clear difference between the weighted and unweighted network, see Fig. 8.12. The inner k-core has few more than 30% of the whole connections, while total sum of the weights remains around 90%. Clearly it makes sense that the inner k-core has most of the information flow, since the prominent and higher weight connections are also in the inner subgraph.

It can also be studied on the obtained layouts the correlation between the degree of the nodes and the k-core. Both quantities are centrality measures and obtaining correlations between them is a very important feature characterizing a network topology. The nodes displayed in the most internal shells (see definition on section 3.2.6) are those forming the central core of the network, the degree-coreness correlations corresponds to the fact that the prominent nodes are most likely high-degree hubs of the network. This observation is indeed obtained in many real-world networks with a clear hierarchical structure, as the Internet at the Autonomous System level or the air transportation network (defined on section 3.4.3).

By other hand, the presence of hubs in external shells is typical of networks without a clear global hierarchical structure as the world-wide-web or the Internet Router Level (defined on section 3.4.1). An emerging of a star-like configurations appears with high degree nodes connecting only to very low degree nodes. These nodes are rapidly excluded of the k-core decomposition, which can lead that high degree nodes local hub in the external k-shells or being excluded from the inner k-cores.

In the world tourism network the countries with high degree, or hubs, also belong to the inner k-core, see Fig. 8.13. This is similar to the air transportation network and the Internet, evidencing a global hierarchical structure with the main tourist destinations as hubs that are also connected to each others. It forms a strong central network, with most of the information flow, also most traffic (tourist arrivals) and highly interconnected, concentrating the main functionality of the network.

8.6 Conclusions

On the world tourism network degree-degree correlations (or assortativity) show disassortative mixing, where high-degree countries, hubs - do play an important role on connecting peripheral clusters of countries attaching themselves to low-degree ones. It is quite notable that on tourism, presently representing the largest set of humans traveling abroad is observed a dissasortative behavior, typically characterizing economic and technological networks.

This is contradicting the review that tourism is a sector with social network behavior. Important questions prompt up, how do economic and technological bases over social ones influence tourism stakeholders? How are economic ties related with consumer behavior - the tourists?

Alternatively, the degree-degree correlations considering weights, measure the affinity to connect with high or low degree countries according to flow dimension. On the tourism network the high degree countries have larger tourist flows between them, as expected since their ties carry higher flows. However, small degree countries have their larger flows with other low degree nodes.

Countries with intermediate values of degree do not have defined affinities showing a dispersal of the magnitude and degrees of their neighbors. The comparison of both measures is displayed in Fig. 8.2. Clustering, also known as transitivity or presence of high number of triangles [Watts and Strogatz, 1998], depicts the structure of the traveling network. The variations of clustering coefficient definition, for (un)weighted and (un)directed networks, relates the transitive triplets with the way countries acquaintance with one another.

The clustering reveals that the traveling acquaintances can be suited by a generalized random network, in accordance with previous results [Miguéns and Mendes, 2008b]. The variations on the clustering depict the evolution dynamics affected by global economic factors. The 2001 tourism slowed down due to September 11's, and in 2003 the biggest slowed down in tourism affected by Iraq war, SARS and the prevailing weak economy, implied an increase of the clustering coefficient that may be related with a lose of competitiveness.

Traveling is usually related with the motives of humans to go to an environment other their usual. Interestingly, the travelers network does not reveal a social acquaintances behavior, but rather a nonsocial one. In this way traveling has similarities to economic and technologic networks. The way the weights are distributed on the triplets links, show a rich club phenomenon [Zhou and Mondragón, 2004] and tendency to agglomerate flow on interconnected groups with higher weight.

On a weighted network the triplets can be more or less connected depending on their links homogeneity. To understand the high heterogeneity of weights on a triplet one introduced two clustering coefficients. One to know how well connected are the neighbors on a triplet, the other depicting adjacent links significance to the relation. Hubs have dominant adjacent links, with weakly connected neighbors, in this way hubs connected otherwise disconnected regions. Oppositely low degree countries have significant strongly connected neighbors and lower weights between themselves and the neighbors.

In this way we show that besides transitivity, clustering relates to other phenomenons (rich club, competitiveness) and depicts acquaintances structure.

Networked Tourism: Case Studies 9

Introduction

The increasing availability of real network data and the current capacity to be able to analyze large data sets has enhanced analytical methods to characterize networks, extending our knowledge on the description of these systems. Networks are a fashionable topic on many real-world systems, such as airline connections, financial relations, companies, partnerships, ecological networks, movies actors, world trade, citations, metabolic, neural, world-wide-web, food webs, email networks, human acquaintance patterns, among others [Dorogovtsev and Mendes, 2003; Strogatz, 2001]. The network approach facilitates a holistic, rather than a focalized, perspective on the destination.

Human mobility and networks have been studied during the last years on a regional and national scale. The results and discussed on section 9.1, on whether there are common patterns on a regional and national with our international tourism network, of human travel mobility.

Around 46% of international tourism has air traffic as mean of transportation, and significantly. Although the configuration of the air transportation network and the tourism network reveal a structural dissimilarity, see section 9.2. A question arises on whether it can be influenced by political national organization and also the way charters can be the solution to overcome the structural dissimilarity of the way airports and tourist destinations are organized. During the last decade some tourism systems were studied using network analysis, the overall pattern of those nets is compared on section 9.3.

9.1 Human Mobility: regional, national and international scale

The emerging information society is widely expected to experience massive embedding of both fixed and portable devices into our local physical spaces, with more and more devices having the capacity to initiate, store and communicate information and content in all aspects of life. This results in significant challenges for communication and information provision, based on required scalability, heterogeneity, re-configurability and dynamicity.

Tourism seeks to embed in devices as the key characteristics that enable tourists to adapt and exhibit agility. As having on the moment access to all about the destination. The social network of tourists based on a relations between fixed and mobile devices, as defined by the human user and their particular actions and behavior with respect to each other and technology in the environment, are expected to bring new way of communication between the tourist and its usual environment, as the tourist and the destination. Social networks are intimately connected with the human and are the basis for the social networks paradigm. To understand this new form of interacting are expected to anthropologically change the role of the tourist as a information and experience broker between cultures. Relations between devices can be inherently flexible, based on casual interactions, using social models of trust and security, and without the need for ”always-on” connectivity. Furthermore, such relations build into social networks with desirable and inclusive properties that can be exploited for communication and knowledge acquisition for large numbers of devices.

Brockmann et al. [2006] reported a quantitative assessment of human travel on geographical scales by investigating the circulation of individual dollar bills in the United States. The analysis was based on data collected at the Internet bill-tracking game www.wheresgeorge.com. The idea of the game is simple: a large number of banknotes is marked by a player and brought

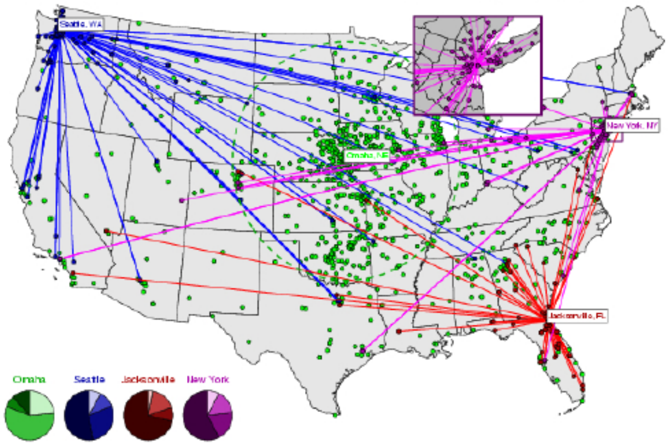


Fig. 9.1: Short time trajectories of individual dollar bills. Lines connect the initial entry location and the location where the bill was reported less than a week after inital entry. Source: [Brockmann et al., 2006].

into circulation. If another person receives a marked banknote, she can register its current location online and return the bill into circulation.

Since wheresgeorge.com started more than 50 million dollar bills have been registered and approximately 10 million have been reported again. Based on a dataset of over a million individual displacements we found that the dispersal of dollar bills is anomalous in two ways. First, the distribution of traveling distances decays as a power-law, indicating that the movement of individuals is reminiscent of superdiffusive, scale free random walks known as Lévy flights. However, computing the time for an initially localized ensemble of dollar bills to redistribute equally within the United States, we found that this time is much longer than predicted by the simple Lévy flight picture.

A deeper analysis of the temporal aspects of the dataset showed that the probability of remaining in a small, spatially confined region for a time T is dominated by algebraically long tails as well. This property, which typically yields subdiffusive dispersal competes with the superdiffusive impact of long jumps and attenuates superdiffusive dispersal. It was shown that the dispersal is an ambivalent effectively superdiffusive process.

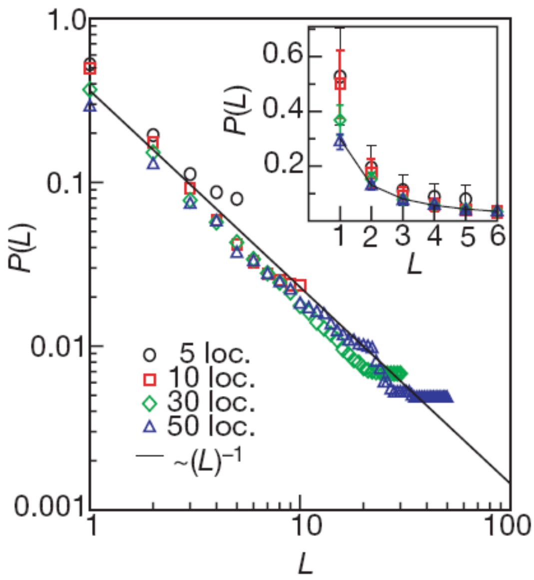


Fig. 9.2: Mobility patterns and mobile phones. A Zipf plot showing the frequency of visiting different locations. Source: [Gonzalez et al., 2008].

Each consecutive sighting of a bank note reflects the composite motion of two or more individuals who owned the bill between two reported sightings. Thus, it is not clear whether the observed distribution reflects the motion of individual users or some previously unknown convolution between population-based heterogeneities and individual human trajectories.

On a national scale and contrary to bank notes, mobile phones are carried by the same individual during the daily routine, offering a good approximation to capture individual human trajectories, see Fig. 9.1. Gonzalez et al. [2008] by analyzing a data set to explore the mobility pattern of individuals, consisting of the mobility patterns recorded over a six-month period for $100,000$ individuals selected randomly from a sample of more than 6 million anonymized mobile phone users.

It was found by Gonzalez et al. [2008] that in contrast with the random trajectories predicted by the prevailing Lévy flight and random walk models [Brockmann et al., 2006], human trajectories show a high degree of temporal and spatial regularity, each individual being characterized by a time independent characteristic travel distance and a significant probability to return to a few highly frequented locations. Moreover the individual travel patterns are described by a single

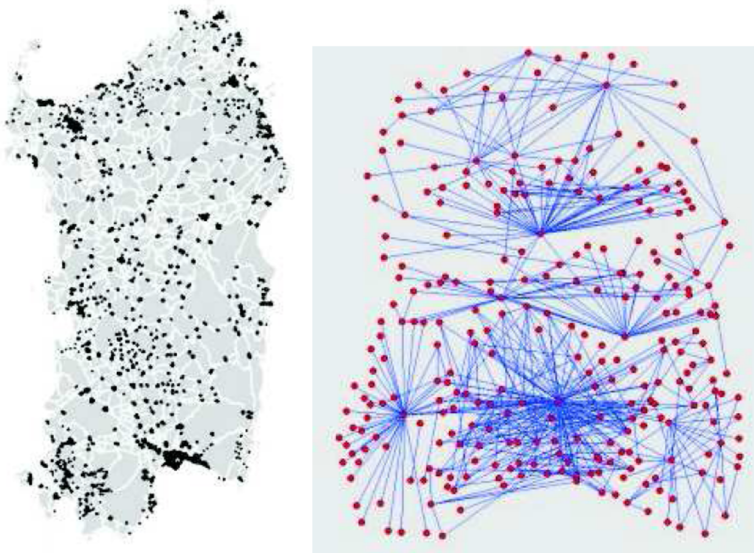


Fig. 9.3: Interurban network on Sardini. Source: [de Montis et al., 2007].

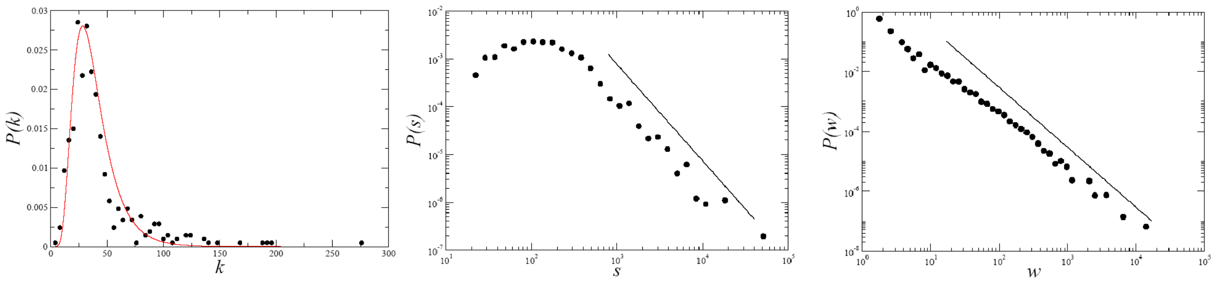


Fig. 9.4: The degree, weight and strength distributions of the interurban traffic. Source: [de Montis et al., 2007].

spatial probability distribution, indicating that, despite the diversity of their travel history, humans follow simple reproducible patterns. This finding are taken to as reproducing travel patterns that can impact all phenomena driven by human mobility, from epidemic prevention to emergency response, urban planning and agent-based modeling.

On a regional scale de Montis et al. [2007] studied the interurban traffic on Sardinia region, Italy, with 375 municipalities and $1,600,000$ inhabitants. It compromised a weighted network representation and quantitative characterization of the capacity of attraction of each urban center on workers and students, where nodes correspond to towns and the edges to the volume of people commuting among those, see Fig. 9.3.

The configuration of the interurban traffic was represented by a small-world random [de Montis et al., 2007], see Fig. 9.4. Like our studied network of the world tourism, also on the interurban traffic a weighted representation of the network allowed by the inclusion of traffic data complements the resulting picture with important information relating the commuter traffic to the available topological connectivity. de Montis et al. [2007] found that both the weights and the strengths are very broadly distributed, confirming the necessity of including those measures in a realistic description of the representation.

Moreover a scale-free behavior was found for the weights and strength (commuter traffic handled by the municipality) distribution. See the distribution of the connectivity, weight and strength on Fig. 9.4.

9.2 Air Transportation, Social Networks and Tourism

Transport is an important and essential linkage between tourist origin and destination. Historically transport has played a vital role in the development of tourism, revolutionized by the development of the railway in the nineteenth century and the private car in the second half of the twentieth century. Air transportation is the dominant mode of international movements of passengers, see Fig. 8.8, and has greatly contributed to reduce distances. Technology developments (see section 2) has also significantly extended the range of aircraft, so that while 40 years ago aircraft were just

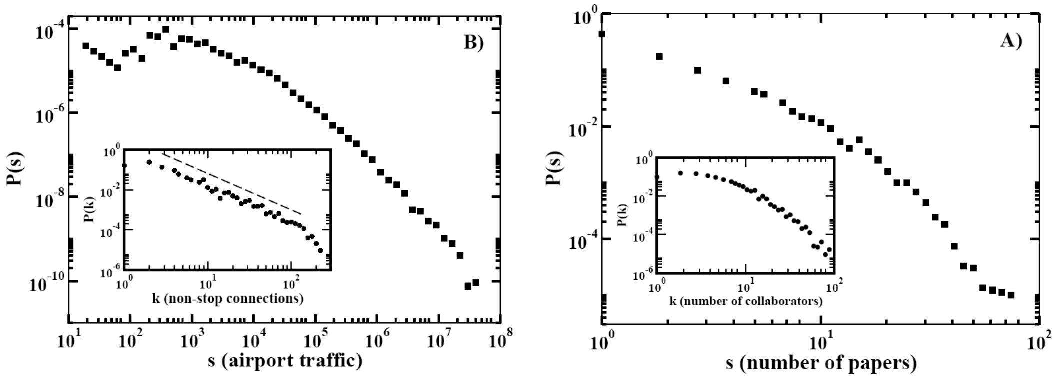


Fig. 9.5: Strength distributions of airports traffic and number of papers per author. Source: Barrat et al. [2004a]

Tab. 9.1: Airports network and world tourism network.

airports	world tourism network
dissassortative	assortative
directed	undirected
random, scale-free	scale-free, scale-free height

beginning to be capable of crossing the Atlantic without stopping at intermediate places they are now capable of making trips of up to hours duration, and the introduction of the jet engine also considerably reduced distances.

Rising affluence of air transportation the growth of air transport is highly correlated to income and economic output growth, as the population of developed countries became more affluent, a greater amount of disposable income became available for leisure. Lower Airfares decrease constantly as a consequence of technical improvements and growing demands making air transportation affordable to the general public, with the low-cost airlines changing the market competition.

On section 3.4.3 is described the air transportation network. In general the network shows also a scale-free behavior when considering the traffic. Contrary to our results of the world tourism network the air transportation network has a scale-free bahavior on the connectivity.

Contrary to the airports network, the world tourism network is highly directed. A portuguese tourist taking holiday in Greece, counts for the world tourism network has a directed connection from Portugal to Greece. A greek tourists taking holidays in Portugal is counting as a one directed connection from Greece to Portugal. To the air transportation network the some tourist most probably takes a flight over another country, to travel through a central airport. Imagining that both flight go over Frankfurt. Then, to the airports networks two connections are added, Portugal to Germany and Greece to Germany. The airports network is not directed, because flights are counted on number of seats available between cities, see Fig. 9.6.

The structural difference just described is also related with the dissassortativity of the world tourism network and the assortativity of the airports network, see Tab. 9.1. While the airports are connected in a way that high degree airports connect to other high degree airports, on the tourism destinations the high degree destinations (more popular) play a central role on connecting to the low degree destinations (less popular). We may question how can two networks so close related with one another have such a strong structural difference. The reasons can be related with the high cost of airports and strong political influence on airports construction.

Along with that it can be that tourism overcome the assortativity of the airports connectivity by introducing charters. This way it is avoided that a tourist goes to a central airport, or hub. On the tourism literature just a few studies are on the topic of charters [Jergensen and Solvoll, 1996; Gillmor, 1996; Mossberg, 1995]. While Jergensen and Solvoll [1996] developed a demand model and Gillmor [1996] a studied the changes on the outbound tourism in Ireland due to the charters, Mossberg [1995] analyses the customer satisfaction. Therefore, no literature review was encounter that could possible relate with our findings.

9.3 The Networked Webspace and Tourism

The internet besides providing an easy search for booking services on a tourism destination, is also an easy channel to find information about a destination.

On the world-wide-web each page contains information and accessing that information without a search engine can be a difficult task. The function of a web search engine is to retrieve information on the majority of the webpages and to categorize them according to their information. The approach of an online destination can consider the hyperlinks between Web sites, or even links between pages in Web sites. The use of search engines to categorize information about a destination directly influences its online image.

A search (on Google and Yahoo!) was performed to obtain a list of the fifty Web sites using the keyword ”Tourism Lisbon”. The first fifty unique sites on tourism attractions, were selected alternatively from both search engines. On each Web site, Lisbon tourism attractions that appear on at least 10% of the Web site were selected. The study was geographically limited to the city of Lisbon, Portugal.

This study explores how information about tourism attractions is hard to find and how different information can be available on different Web sites. Network theory is specifically adapted for this study because it focuses on relational systems, where relation between nodes prevail the individual characteristics of nodes. The network of the tourism destination, e-Lisbon, is represented as a bipartite graph, whose nodes can be divided into two disjoint sets, on the following way. One set is the webpages and the other set is the tourism attractions. In this way, each Web site from the first set connects to the tourism attractions from the other set (nodes) - forming a network where all the nodes are connected to each other (complete graph). We defined that two attractions were connected if they both had information in the same Web site. The same analysis was performed on all Web sites.

Considering that more than one Web site can relate to two attractions. The connection between two attractions on the network representation is labeled with the number of times (weight) they are in a common Web site. Therefore, each pair of tourism attractions, i and j, are weighted link, w_{ij} (see Fig. 9.7). Many real networks also weighted networks. In the case of social networks [Wasserman et al., 1994] it is often relevant to assign a weight (strength) to each edge, measuring how good or strong is a relationship.

A network is assortative (disassortative) with regards to a certain property if it is observed a positive (negative) correlation in that property when considering adjacent nodes. Meaning that a network is assortative if the vertices tend to connect to other vertices which have similar (dissimilar) properties. Mathematically, see Newman [2003a]:

$$r = \frac{M^{-1}\sum_{\phi}(\prod_{i\in F(\phi)} k_i) - [\frac{M^{-1}}{2}\sum_{\phi}(\sum_{i\in F(\phi)} k_i)]}{\frac{M^{-1}}{2}\sum_{\phi}(\sum_{i\in F(\phi)} k_i^2) - [\frac{M^{-1}}{2}\sum_{\phi}(\sum_{i\in F(\phi)} k_i)]}, \tag{9.1}$$

on our case $r = -0.36$, where $F(\phi)$ denotes the set of the two vertices connected by the th link and M is the total number of edges in the network. This measure r is the Pearson coefficient and is positive (negative) for assortative (disassortative) networks ($r = 0$ for a random graph) [Newman, 2003a]. Empirical studies reveal that technological and biological networks appear to be disassortative with respect to the degree, while social networks are generally assortative [Newman, 2003a].

For weighted networks the assortativity coefficient, r^w follows the same meaning, being positive for weighted assortative networks, and negative for weighted disassortative networks. In our case $r^w = -0.4$. The weighted correlation is also assortativity, confirming the topological measure. Assortative networks are resilient to simple target attack, like disease propagation, social networks are more vulnerable than technological and biological networks against attacks or propagations.

Moreover if $r^w > r$ ($r^w < r$), two similar degree nodes are tend to be connected by a higher-weight (lower-weight) edge. Mathematically,

$$r^w = \frac{H^{-1}\sum_{\phi}(w_\phi\prod_{i\in F(\phi)} k_i) - [\frac{H^{-1}}{2}\sum_{\phi}(w_\phi\sum_{i\in F(\phi)} k_i)]}{\frac{H^{-1}}{2}\sum_{\phi}(w_\phi\sum_{i\in F(\phi)} k_i^2) - [\frac{H^{-1}}{2}\sum_{\phi}(w_\phi\sum_{i\in F(\phi)} k_i)]}, \tag{9.2}$$

resulting that similar degrees tend to be connected by higher-weight, high (low) degree nodes tend to be strongly connected to high (low) degree nodes.

The webspace of a tourism destination was represented by Baggio [2007], on the island of Elba, coast of Tuscany, Italy, where the vertices are companies and the links between them are hyperlink connections. Baggio [2007] also performed obtaned an dissassortative behavior between companies and also obtained a similar pattern on formal partnerships and hyperlinks on the webspace.

In general, like the world tourism network (see section 8), also both the network of the sightseeings (see Fig. 9.7 (a)) on a web search and the network of the webspace of a tourist destination (see Fig. 9.7 (b)) are dissassortative. Therefore the same behavior, where hubs play a key role on connecting to perpherical nodes is observed on a variety of tourism networks, and can be a general behavior of tourism networked systems.

9.4 Patterns on Multidestination

In the last decade the concept of networks has been growing in attention within the tourism literature, on a long type of application, from multidestination, to webspace relations and also networks formed by organizations [Scott et al., 2008b; Shih, 2006; Hwang et al., 2006].

Further Hwang et al. [2006] and Shih [2006] conducted studies to show that travel patterns can be understood as networks, by assessing the structural properties of travel within and between different destinations. A multidestination can be a set of destinations that a tourist visit in a single trip. Shih [2006] studied the Nantou county in Taiwan, a tourist destination (county) which englobe several tourist attractions situated in different (sub)destinations. Among 16 destinations the outcomes resulted on a the understanding of the relational dependency and importance of each destination on the set of destination.

Additionally Zach et al. [2008] studied how visited places in Northern Indiana by tourists show the effect of ”long tail” (see section 3), relating the finding with the Zipf's law and the preferential attachment [Albert and Barabási, 1999].

More than single applications of network theory, these studies bring together common patterns and the empirical evidences that the structure of the network plays an important role on the system performance. And therefore we may call it the networked tourism.

9.5 Conclusions

We have shown that a more complete view of complex networks is provided by the study of the interactions defining the links of these systems.

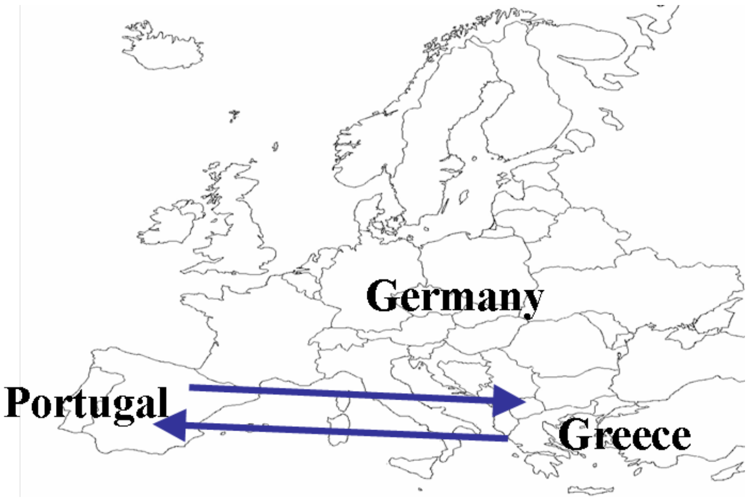
- **Scale-free versus Random Network:** The power-law tail indicates that the probability of having countries with a large inbound and outbound tourism is significant, as the network intensity is dominated by countries with numerous arrivals of tourists [Miguéns and Mendes, 2008b]. Parallel to our results, the bank notes dispersal [Brockmann et al., 2006] describing long range human travel patterns, showing a representation on a spatial and temporal bases. Analogous to our work the study of inter-urban traffic revealed a scale-free behavior when considering traffic rather than connectivity. One might consider a class of networks, highly heterogenic networks, for which new connections have a random growth, but where the highly heterogeneity of flows favor the increase of larger connections. It seems likely that other real-world networks formally showing an exponential distribution, have indeed a power-law distribution on their weights, with some sort of preferential growth. In summary, these results suggest that, the observed on worldwide tourism, is applying to a more widely system, the global human travel pattern;

- **Air Transportation network and the World Tourism Network:** Lew and McKercher [2006] document that the spatial structure of travel can have strong influence on transportation planning, tourism development, and impact management. Although tourism and air transportation are strongly related, as 46% of the international tourism uses air transportation, they have important structural differences. The world tourism network (WTN) is a strongly directed network, while the air transportaion network (AT) is undirected. Moreover, probably the most significant difference is that the ATN is assortative, meaning that the central airports connect with each others and their own subgroup. While the WTN is dissassortative, meaning that most popular tourism detsination play an important role on connecting the less popular destinations. This structural difference and considering the importance of air transportation on tourism, probably the charters are the way to overcome the difference.

- **Global Dissassortativity on the Webspace of Tourism:** in general, like the world tourism network (see section 8), also both the network of the sightseeings (see Fig. 9.7 (a)) on a web search and the network of the webspace of a tourist destination (see Fig. 9.7 (b)) are dissassortative. Therefore the same behavior, where hubs play a key role on connecting to perpherical nodes is observed on a variety of tourism networks, and can be a general behavior of tourism networked systems;

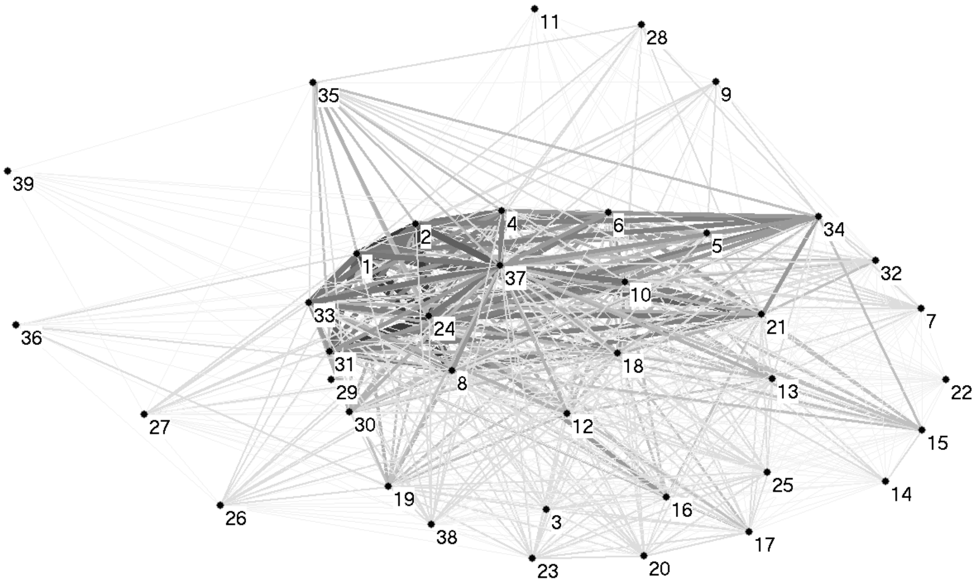
- In the last decade the concept of networks has been growing in attention within the tourism literature, on a long type of application, from multidestination, to webspace relations and also networks formed by organizations [Scott et al., 2008b; Shih, 2006; Hwang et al., 2006]. More than single applications of network theory, these studies bring together common patterns and the empirical evidences that the structure of the network plays an important role on the system performance. And therefore we may call it the networked tourism.

(a) The airports network.

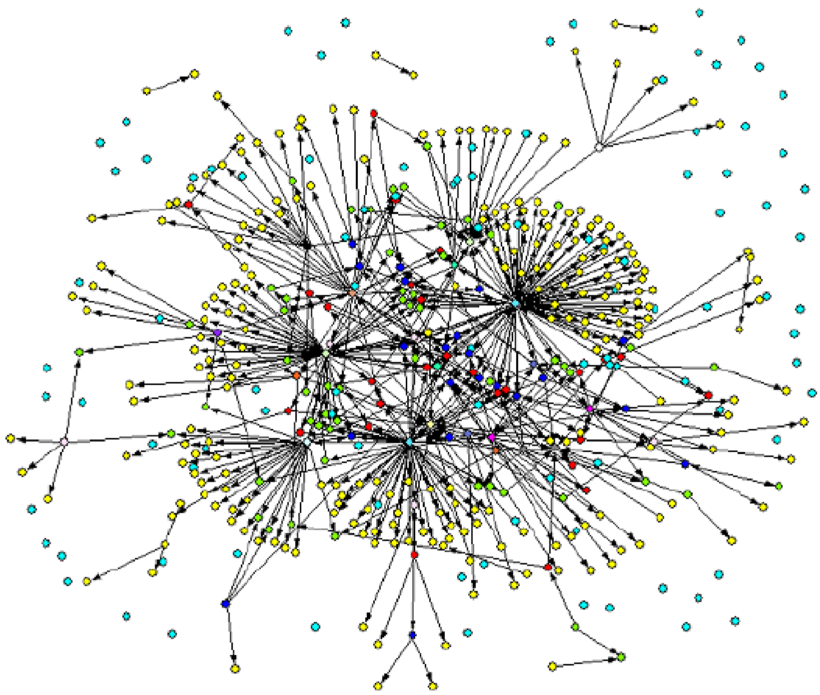


(b) The world tourism network.

Fig. 9.6: Contrary to the airports network, the world tourism network is highly directed. A tourist from Portugal taking holiday in Greece is counting a one directed connection from Portugal to Greece. A greek tourists taking holidays in Portugal is counting as a one directed connection from Greece to Portugal. By other had these tourists most probably take a flight to travel that is going over another country, to a central airport. Imagining that both flight go over Frankfurt. Then, to the airports networks two connections are added, Portugal to Germany and Greece to Germany. The airports network is not directed, because flights are counted on number of seats available between cities.



(a) The e-destination structure of sightseeing attractions in the city of Lisbon. Source: [Miguéns and Corfu, 2008]



(b) E-destination linkages visualization. Source: [Baggio, 2007]

Fig. 9.7: Network visualization of an online tourist destination.

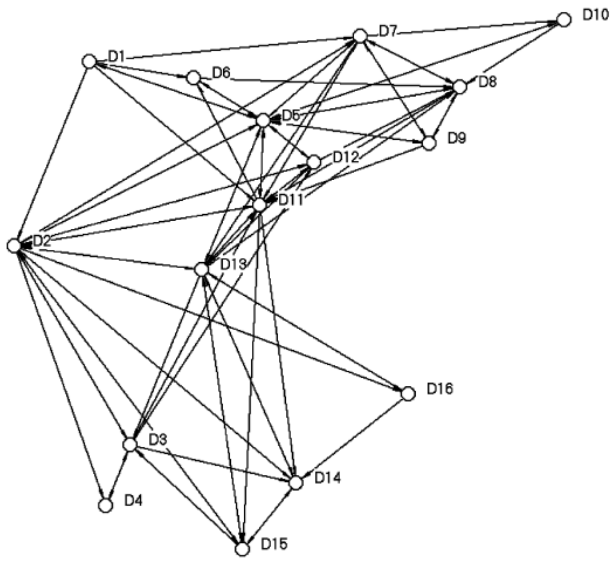
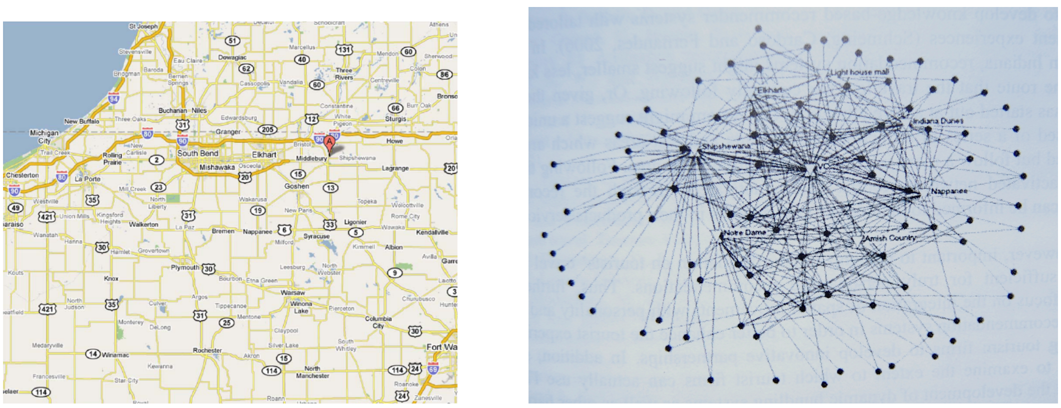


Fig. 9.8: The network of destinations on Taiwan as a tourist destination. Source: [Shih, 2006].



(a) The map of the tourist region, northern Indiana, USA. Source: Google Maps.

(b) E-destination linkages visualization. Source: [Zach et al., 2008]

Fig. 9.9: Network visualization of the tourist attractions on Northern Indiana. Source: [Zach et al., 2008].

10 Conclusions

Networked tourism is a paradigm that reconsiders the way we think about travel and tourism systems. This thesis aimed to develop and adapt adequate methodologies and theories that accomplish present structural challenges of the tourism sector, as well as to study the structural properties of the world tourism network of arrivals and departures, representing the largest data of human travel outside their usual environment.

The way tourists travel between countries is a large network that is shaping our society, changing behavior, influencing economies and partly defining a new era. More than witnessing a growing rate of tourists on a world scale, this thesis brings answers to some important question on the travel and tourism industry. How do tourist destinations group themselves? How do travelers choose their destination? How do macro-economic factors influence the destinations neighborhood? What is the importance and role of hubs (most visited destination) on a world scale? These are some of the questions that this thesis approaches by using network theory.

Networks do not belong to a single discipline, but are rather a part of many fields of knowledge, from mathematics, physics, management, economics and sociology. The simple fact that people interact on communities, organizations compete among themselves, web pages are linked in a large net of information, stocks are correlated, and tourists interact to decide for a destination, show that relationships or connections are important, and essential. On this sense the research accomplished a literature review on the different strands of network theory. The vast review, from graph theory to complex networks and social network analysis was the starting point of the research. The common measurements, and developments over the last decade on complex networks, are presented as well as applied to the tourism (data from WTO), bringing new methodologies into the field.

The network studied is the world network of the tourist arrivals and departures, from each country to another. In this sense it is known how many tourists are going annually from Portugal to Spain, but also how many are going from Spain to Portugal. So what ever is the mean of transportation, or the reason of the travel, the network has the absolute number of all the tourists traveling between every two countries in the world. The described net has special features. For example, the number of tourists going in one direction is in most of the cases considerable different from the other direction, like between Portugal and Spain. The network is therefore called a directed network. Another property is that flows have different scales. The inbound tourism of a given country can differ to others on scale. For example, the highest number of tourist arrivals (on the year 2004) between every pair of countries was from United States of America to Mexico, with $19.369.677$ tourists, but the average value is 81.813 tourists, revealing a high heterogeneity of links.

An important statistical property to directed networks is the reciprocity, which on the tourism network means the appetency to exchange tourists (see section 6.1 and 6.4). A percentage of 60% of countries are not connected to each other, meaning they do not exchange tourists at all. If one country has tourists visiting from another country, there is only a 25% probability that the second country will have tourists visiting from the first country. And, despite the large number of travelers, the overall connectivity of the network is lower than might be expected. This result means that there are considerable potentialities of grow for a tourism destination. While the global grow of flows allows the already existing tourist origins to increase their flows, it is observed that many of the possible connection do not exist which can be new potential tourist origins.

Another interesting feature of the world tourism network, is that it behaves more like an economic network rather than a social network, see section 8. This observation is obtained on the analysis of the tourism networks (UNWTO) and brings a completely new result when analyzing patterns on the travel and tourism industry. In the tourist network, countries with a high degree (those that are popular tourist destinations) are more likely to be neighbors with countries with a low degree (less popular tourist destinations). The most popular tourist destinations, called hubs, play a central role on connecting peripheral destinations. This behavior is similar to economic and transportation networks, which have patterns where popular central hubs have many inbound connections from peripheral nodes. Further, this finding could question the common notion of culture as the driving force of tourism.

On the overall, the world tourism network is scale-free (see section 7.4) over four orders of magnitude. It describes international short traveling range to long travels, on a global scale. The probability distribution that a connection has a certain volume displays a power-law decay. It is worth noticing that scale free networks have the ability to change scale in order to meet any level of demand. So as travelers increase, as expected, it is the countries that are already the popular tourist destinations that are more likely to add new connections (receive highest percentage of new tourists).

The network changes in scale, but does not change in the relation of connections, even as the number of travelers dramatically increases or decreases. The scale-free versus random network means that when comparing two countries, a popular destination and a less popular destination, their tourists can come from approximately the same number of countries, what distinguishes them is the volume, the number of tourists per connection. It is comparable to highways on a country, most of them have the same number of connections with other roads, although the amount of vehicles per road differs significantly. This observation can have impact when deciding on national strategies. Considering the growing evolution of tourism and its patterns, is it better to diverse the number of tourism origin countries visiting our destination, or is it better to increase the traffic of the already existing connections?

When analyzing this network some properties emerge that makes it unique and with particular implications. For example, the simple fact that the network is a scale-free is related with the possible evolution process of the net. This is related with the Zipf 's and Pareto law, this results show that these laws also apply to the world tourist arrivals and departures network.

Similarities with other real-world networks were analyzed. Tourism, a sector highly dependent on other economic and society systems, shows common patterns and some structural differences with the air transportation network. While in the air transportation central airports (hubs) connect to hubs, showing an assortative behavior, the tourism destinations network show a dissassortative behavior, where hubs play an important role on connecting otherwise disconnected destinations. This can be explained by political reasons and can also influence the growth of the number of flights performed by charters (see section 9.2).

A self-similarity with other studied tourism networks shows the recent interest on this topic among researchers on tourism. For example, a net of organizations linked on the webspace, also show a dissassortative behavior, and is also a scale-free network. Other measurements typical of network theory are recently used on tourism literature, like betweeness, k-core and clustering. Assuming the importance of relationships to tourism systems, the behavior found can influence the way new linkages are made between organizations, being it partnerships or common marketing strategies. Tourist destinations are also a networked system, where sightseeing attractions can be considered as nodes and the connections the access between them. Previous studies with this approach also show how central nodes can play an important role on connecting less visited attractions [Shih, 2006].

Moreover the results propose new indicators for tourism. The macroeconomic factors that affect tourism can be monitored by using the clustering coefficient. On the world net of tourist arrivals and departures, is observed a clustering coefficient that relates with economic and natural factors over the years. In general, on the years with a decrease of global tourist arrivals, for example on the year of 2001 due to the September 11^{th}, the clustering coefficient also changes, reflecting a loose of long-haul connection on the international tourism. This result can be used to understand which countries and regions loose competitiveness during a certain period, by a quantification of connectivity among the neighbors of a country. Therefore, monitoring these patterns can be a rich tool for governments and world tourism organizations.

The research also shows that the networks as a theoretical framework, bring a new interdisciplinary area into the tourism academia (see section 5.1). This should open a new recognition of work done in other disciplines, even if it lies outside our normal purview, to improve our understanding of the tourism systems. Previous research on international movements of tourists and travelers only conducts qualitative analysis and quantitative methods based on individual properties from international tourists and undertakes no modeling of interconnected system as described in this thesis. The application of network theory appears to have merit, to enhance strategist' and managers understanding of behavior among tourists linking micro and macro level analysis.

Network analysis allows visualizing structural differences that cannot be accomplished with traditional methods. Recent developments in network analysis approaches, such as simulation software and programs, to test multiple theories at multiple analysis levels [Batagelj and Mrvar, 2002; Borgatti et al., 2002] promise to further enhance the understanding of nets in travel and tourism industry. On a tourism research point of view, the understanding of the evolving structure of travel using a network analysis perspective offers relevant implications, beyond destinations structure and tourists behavior modeling.

To our knowledge, this is the first empirical evidence on a worldwide scale observing human travel dynamics on tourism flows.

References

Agar, M. (1996). *The Professional Stranger : An Informal Introduction to Ethnography*. Academic Press. pages 102

Ahasa, R., Aasaa, V., Roosea, A., Markb, V., and Silma, S. (2008). Evaluating passive mobilenext term positioning data for tourism surveys: An estonian case study. *Tourism Management*, 29(3):469–486. pages 91

Albert, R. and Barabási, A.-L. (1999). Emergence of scaling in random networks. *Science*, 286:509–512. pages 45, 50, 51, 54, 67, 73, 97, 104, 143, 185

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97. pages 44, 48, 49, 50, 109, 143

Albert, R., Jeong, H., and Barabasi, A.-L. (1999). Internet: Diameter of the world-wide-web. *Nature*, 401(6749):130–131. pages 45, 68, 73, 97

Alvarez-Hamelin, J., Dall'Asta, L., Barrat, A., and Vespignani, A. (2006a). k-core decomposition: a tool for the visualization of large scale networks. *CANADA*, page 41. pages 61

Alvarez-Hamelin, J. I., Dall'Asta, L., Barrat, A., and Vespignani, A. (2006b). Large scale networks fingerprinting and visualization using the k-core decomposition. In Weiss, Y., Schölkopf, B., and Platt, J., editors, *Advances in Neural Information Processing Systems 18*, pages 41–50. MIT Press. pages 61

Amaral, L. A. N., Scala, A., Barthélémy, A., and Stanley, H. E. (2000). Classes of small-world networks. *Proc. Natl. Acad. Sci. U.S.A.*, 97:11149. pages 48, 73, 164

Araujo, L. and Bramwell, B. (2002). Partnership and regional tourism in brazil. *Annals of Tourism Research*, 29:1138–1164. pages 87

Arenas, A., Danon, L., Díaz-Guilera, A., Gleiser, P., and Guimerà, R. (2004). Community analysis in social networks. *Eur. Phys. J. B*, 38:373–380. pages 76

Axelsson, B. and Easton, G. (1992). *Industrial Networks: A New View of Reality*. Routledge, London. pages 85

Baggio, R. (2007). The web graph of a tourism system. *Physica A*, 379(2):727–734. pages 91, 93, 94, 97, 184, 189

Baht, S. S. and Milne, S. (2007). The impact of networks on cooperation in destination marketing: The case of www.purenz.com. In L. Luisa Andreu, J. Gnoth, . M. K., editor, *Proceedings of the 2007 advances in tourism marketing conference [CD-ROM]*, Valencia, Spain. University of Valencia. pages 83, 96

Barabasi, A.-L. (2003). *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. Plume Books. pages 23, 48, 113

Barrat, A., Barthelemy, M., Pastor-Satorras, R., and Vespignani, A. (2004a). The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. USA*, 101:3747. pages 18, 54, 77, 78, 79, 80, 147, 155, 159, 163, 181

Barrat, A., Barthelemy, M., and Vespignani, A. (2004b). Modeling the evolution of weighted networks. *Physical Review E*, 70:066149. pages 18, 70

Barrat, A. and Weigt, M. (2000). On the properties of small-world network models. *Europ.Phys. J.B*, 13:547. pages 63, 66, 159

Barthélemy, M. and Amaral, L. (1999). Small-world networks: Evidence for a crossover picture. *Phys. Lett. A*, 82:3180–3183. pages 66

Batagelj, V. and Mrvar, A. (2002). *Pajek - Analysis and Visualization of Large Networks*, volume 2265. pages 94, 105, 107, 112, 195

Baum, J. A. and Oliver, C. (1996). Toward an intitutional ecology of organizational founding. *Academy of Management Journal*, 39(5):1378–1427. pages 95

Bèdard, F., Louillet, M. C., and Joly, M.-C. (2008). Implementation of a destination management system interface in tourist information centers and its impact. In Höpken, W. and eds., U. G., editors, *Information and Communication Technologies in Tourism 2008*, pages 220–231, Innsbruck. SpringerWienNewYork. pages 35

Beltran, J. (2000). Indigenous and traditional peoples and protected areas: Principles, guidelines and case studies. Technical report, IUCN, Gland, Switzerland and Cambridge, UK. pages 89

Bender, E. A. and Canfield, E. R. (1978). The asymptotic number of labeled graphs with given degree sequences. *J. Combin. Theory A*, 26:296–307. pages 64

Benford, F. (1938). The law of anomalous numbers. *Proc. of the American Philosophical Society*, 78:551–572. pages 18, 133, 134

Bhat, S. S. and Milne, S. (2008). Network effects on cooperation in destination website development. online. pages 93

Bollabas, B. (1984). The evolution of random graphs. *Trans. Amer. Math. Soc.*, 286:257–274. pages 61, 64

Bollobás, B. (1985). *Random Graphs*. Academic Press, London. pages 63

Borgatti, S. P., Everett, M. G., and Freeman, L. C. (2002). Ucinet 6 for windows: Software for social network analysis. pages 94, 105, 107, 112, 195

Bornholdt, S. and Ebel, H. (2001). World wide web scaling exponent from simon's 1955 model. *Phys. Rev. E*, 64(3):035104. pages 69

Bornholdt, S. and Schuster, H. G. (2003). *Handbook of Graphs and Networks: From the Genome to the internet*. Wiley-VCH. pages 64

Bramwell, B. and Sharman, A. (1999). Collaboration in local tourism policy making. *Annals of Tourism Research*, 26(2):392–415. pages 87

Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Social Networks*, 30(2):136–145. pages 57, 95

Brass, D. J. (1984). Being in the right place: A structural analysis of individual influence in an organization. *Administrative Science Quarterly*, 29:518–539. pages 95, 104

Brass, D. J. and Brurkhardt, M. E. (1993). Potentila power and power use: An investigation of power and behavior. *Academy Management Journal*, 36(3):441–470. pages 95

Breukel, A. and Go, F. M. (2008). Knowledge-based network participation in destination and event marketing: A hospitality scenario analysis perspective. *Tourism Management*, pages 1–10. pages 90

Brockmann, D., Hufnagel, L., and Geisel, T. (2006). The scaling laws of human travel. *Nature*, 439:462. pages 176, 177, 178, 186

Buchanan, M. (2003). *Nexus: Small Worlds and the Groundbreaking Theory of Networks*. W. W. Norton & Company. pages 44, 48, 49, 50, 109

Buhalis, D. (1998). Strategic use of information technology in the tourism industry. *Tourism Management*, 19(5):409–421. pages 35, 36, 37, 39

Buhalis, D. (2000). Marketing the competitive destination of the future. *Tourism Management*, 21(1):97–116. pages 87

Buhalis, D. and Costa, C. (2005). *Tourism Management Dynamics*, volume 1. Elsevier, butterworth heinemann edition. pages 31

Buhalis, D. and Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the internetthe state of etourism research. *Tourism Management*, 29(4):609–623. pages 35

Buhalis, D. and Main, H. (1998). Information technology in peripheral small and medium hospitality enterprises: Strategic analysis and critical factors. *International Journal of Contemporary Hospitality Management*, 10(5):198–202. pages 35, 90, 97

Burt, R. S. (1995). *Structural Holes: The Social Structure of Competition*. Harvard University Press. pages 53, 57, 82, 85, 93, 94, 95, 104

Caldarelli, G., Battiston, S., Garlaschelli, D., and Catanzaro, M. (2004). Emergence of complexity in financial networks. *Ben-Naim E, Frauenfelder H, Toroczkai Z (eds) Lecture Notes in Physics*, 650:399–423. pages 151

Caldarelli, G., Capocci, A., Rios, P. D. L., and Muñoz, M. A. (2002). Scale-free networks from varying vertex intrinsic fitness. *Phys. Rev. Lett.*, 89:258702. pages 67

Cartwright, D. and Harary, F. (1956). Structural balance: a generalization of heider's theory. *Psychological Review*, 63:277–293. pages 75

Castells, M. (1996). *The Information Age: Economy, Society, and Culture.*, volume I of *The Rise of the Network Society*. Oxford. pages 23, 89, 90, 113

Cetron, M. (2001). *Tourism and hospitality in the 21_{st} century*. The world of today and tomorrow: the global view. Butterworth-Heinemann, Oxford. pages 40

Chowell, G., Hyman, J. M., Eubank, S., and Castillo-Chavez, C. (2003). Scaling laws for the movement of people between locations in a large city. *Physical Review E*, 68:066102. pages 163

Chung, F. and Lu, L. (2002). The average distances in random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA*, 99:15879. pages 64, 66, 67

Cohen, R. and Havlin, S. (2003). Scale-free networks are ultrasmall. *Phys. Rev. Lett.*, 90:58701. pages 67

Colloquium, E., editor (1998). *Developing a classification of network typologies*. pages 85

Cooper, C., Fletcher, J., Fyall, A., and Gilbert, D. (2008). *Tourism: Principles and Practice*. Financial Times/ Prentice Hall. pages 27

Costa, C. (1996). *Towards the Improvement of the Efficiency and Effectiveness of Tourism Planning and Development at the Regional Level - Planning and Networks. The case of Portugal.* PhD thesis, University of Surrey, Guildford. pages 82

Costa, C., Breda, Z., Costa, R., and Miguéns, J. (2007). *The benefits of networks to SMEs*, chapter Network Analysis and Tourism: From Theory to Practice, pages 96–114. Aspects of Tourism Book Series. Channel View Publications. pages 87

Cross, R. L., Parker, A., and Borgatti, S. P. (2000). A birds-eye view: Using social network analysis to improve knowledge creation and sharing. *Knowledge Directions*, 2(1):48–61. pages 104

Crotts, J. (April 1998). Antecedents of supplier's commitment to wholesale buyers in the international travel trade. *Tourism Management*, 19:127–134(8). pages 85

Crotts, J. and Aziz, A. (1998). Antecedents of supplier's commitment to wholesale buyers in the international travel trade. *Tourism Management*, 19:127–134. pages 82

Crouch, G. I. (1994). The study of international tourism demand: a review of findings. *Journal of Travel Research*, 33:12–23. pages 39

Curran, J. and Storey, D. (1993). *Small Firms in Urban and Rural Locations*. Routledge, London. pages 83

Davidsen, J., Ebel, H., and Bornholdt, S. (2002). Emergence of a small world from local interactions: Modeling acquaintance networks. *Phys. Rev. Lett.*, 88:128701. pages 65

de Montis, A., Barthelemy, M., Chessa, A., and Vespignani, A. (2007). The structure of interurban traffic: A weighted network analysis. *Environment and Planning B: Planning and Design*, 34(5):905–924. pages 20, 54, 143, 151, 152, 157, 159, 163, 179, 180

Degenne and Forsé (1999). *Introducing Social Networks*. Sage Publications. pages 49

Degenne, A. and Forse, M. (1999). *Introducing*. networks. London: Sage Publications. pages 128

Di-Maggio, P. J. and Powell, W. W. (1983). The iron cage revised: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48:147–160. pages 96

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271. pages 57

Dorogovtsev, S. N., Golsev, A. V., and Mendes, J. F. F. (2006). k-core organization of complex networks. *Phys. Rev. Lett.*, 96:040601. pages 61

Dorogovtsev, S. N. and Mendes, J. F. F. (2003). *Evolution of networks: From biological nets to the internet and WWW*. Oxford Univ. Press. pages 48, 54, 79, 103, 104, 109, 151, 175

Dorogovtsev, S. N. and Mendes, J. F. F. (2004). Minimal models of weighted scale-free networks. pages 70, 72

Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N. (2000). Structure of growing networks: Exact solution of the barabasi–albert's model. *Physical Review Letters*, 85:4633. pages 51, 69

Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N. (2001). Size-dependent degree distribution of a scale-free growing network. *Phys. Rev. E*, 63(6):062101. pages 69

Dorogvotsev, S. N. and Mendes, J. F. F. (2000). Evolution of networks with aging of sites. *Physical Review E*, 62:1842. pages 69

Dredge, D. (2006). Policy networks and the local organisation of tourism. *Tourism Management*, 27:269280. pages 89

Dwyer, L. and Kim, C. W. (2003). Destination competitiveness: a model and indicators. *Current Issues in Tourism*, 6(5):369413. pages 39, 40

Ebel, H. (2002). Scale-free topology of e-mail networks. *Phys. Rev. E*, 66:035103. pages 152, 163

Elliot, J. and Johns, N. (1993). The influence of international tourism trends on the design. *International Journal of Contemporary Hospitality Management*, 5(2):6. pages 40

Erdös, P. and Rényi, A. (1959). On random graphs. *Publicationes Mathematicae*, 6:290. pages 62, 63

Erdös, P. and Rényi, A. (1960). On the evolution of random graphs. *Mat. Kutato Int. Kozl*, 64(5):17–60. pages 51, 62, 151

Fagiolo, G. (2006). Clustering in Complex Directed Networks. *ArXiv Physics e-prints*. pages 151, 168, 169

Faloutsos, M., Faloutsos, P., and Faloutsos, C. (1999). On power-law relationships of the internet topology. In *SIGCOMM '99: Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication*, volume 29, pages 251–262, New York, NY, USA. ACM Press. pages 73, 91

Fesenmaier, D. R., Werthner, H., and Woeber, K. (2006). *Destination Recommendation Systems: Behaviroural Foundations and Applications*. CABI Publishing, Oxford. pages 35

Firenze-Turismo (2008). http://www.firenzeturismo.it. accessed online on February 27^{th}. pages 38

Fogelman-Soulie, F. and Herault, J. (1989). *Neurocomputing-Algorithms, Architectures and Applications*, pages 227–236. Berlin: Springer-Verlag. pages 35, 37

Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social networks (Soc. networks)*, pages 215–239. pages 74, 95, 107

Fyall, A., Leask, A., and Garrod, B. (2001). Scottish visitors attractions: a collaborative future? *International Journal of Tourism Research*, 3:211–228. pages 85

Garlaschelli, D., Battiston, S., Castri, M., Servedio, V. D., and Caldarelli, G. (2005). The scale-free topology of market investments. *Physica A: Statistical and Theoretical Physics*, 350(2-4):491–499. pages 70

Gilbert, E. N. (1959). Random graphs. *Ann. Math. Statist*, 30:1141. pages 62, 63

Gillmor, D. A. (1996). Evolving air-charter tourism patterns: change in outbound from the republic of ireland. *Tourism Management*, 17:9–16. pages 182

Giménez-Gracía, V., Martínez-Parra, J., and Buffa, F. (2007). Improving resource utilization in multi-unit networked organizations: The case of a spanish restaurant chain. *Tourism Management*, 28:262270. pages 81, 90

Gnoth, J. (2002). Leveraging export brands through a tourism brand. *The Journal of Brand Management*, 9(4/5):262–280. pages 87

Goh, K., Kahng, B., and Kim, D. (2001). Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.*, 87:278701. pages 66, 67

Goltsev, A. V., Dorogovtsev, S. N., and Mendes, J. F. F. (2006). k-core (bootstrap) percolation on complex networks: Critical phenomena and nonlocal effects. *Phys. Rev. E*, 73:056101. pages 61

Gonzalez, M. C., Hidalgo, C. A., and Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *Nature*, 453(7196):779–782. pages 178

Granovetter, M. S. (1973). The strength of weak ties. *The American Journal of Sociology*, 78(6):1360–1380. pages 23, 38, 49, 54, 75, 86, 93, 94, 96, 105, 107

Gretzel, U., Hyan-Yoo, K., and Purifoy, M. (2008). Online travel review study: The role and impact of online travel reviews. research, , Laboratory for Intellegent Systems in Tourism, College Station. pages 36, 87

Gössling, S. (2002). Global environment consequences of tourism. *Global Environment Change*, 12(4):283–302. pages 31

Guimera, R., Mossa, S., Turtschi, A., and Amaral, L. A. N. (2005). The worldwide air transportation network: Anomalous centrality, community structure, and cities global roles. *Proc. Natl. Acad. Sci. USA*, 102:7794–7799. pages 45, 79, 142, 151

Hall, C. M. (2000). *Tourism collaboration and partnerships: Politics, practice and sustainability*, chapter Rethinking collaboration and partnership: A public policy perspective., pages 143–158. Channel View Publications, UK. pages 39, 85, 87

Hall, C. M. (2005). Systems of surveillance and control: commentary on 'an analysis of institutional contributors to three major academic tourism journals: 19922001'. *Tourism Management*, 26:653656. pages 18, 110, 111

Halme, M. (2001). Learning for sustainable development in tourism networks. *Business Strategy and the Environment*, 10:100–114. pages 83, 85

Harshaw, H. and Tindall, D. (2005a). Social structure, identities, and values: A network approach to understand people's relationships to forests. *Journal of Leisure Research*, 37(4):426–449. pages 83

Harshaw, H. W. and Tindall, D. B. (2005b). Social structure, identities, and values: A network approach to understanding peoples's relationships to forecast. *Journal of Leisure Research*, 37:426. pages 82

Hawkins, D. (2004). A protected areas ecotourism competitive cluster approach to catalyse biodiversity conservation and economic growth in bulgaria. *Journal of Sustainable Tourism*, 12:3. pages 91

Heider, F. (1946). Attitude and cognitive orientation. *Journal of Psychology*, 21. pages 74, 75

Hogan, C. (2000). Pan parks: A synergy between nature conservation and tourism in europe's protected areas. communications, PAN Parks Communications office. pages 89

Holder, J. S. (1991). Tourism, the world and the caribbean. *Tourism Management*, 12(4):291–300. pages 35

Holland, P. W. and Leinhardt, S. (1970). A method for detecting structure in sociometric data. *American Journal of Sociology*, 70:492–513. pages 76

Holme, P., Park, S. M., Kim, B. J., and Edling, C. R. (2007). Korean university life in a network perspective: Dynamics of a large affiliation network. *Physica A*, 373:821. pages 156

Hunziker, W. and Krapf, K. (1942). Grundrisse der allgemeinen fremdenverkehrslehre. *Polygraphischer Verlag*, 1:392. pages 29, 30

Hwang, Y.-H., Gretzel, U., and Fesenmaier, D. R. (2006). Multicity trip patterns – tourists to the united states. *Annals of Tourism Research*, 33:10571078. pages 92, 95, 97, 185, 187

IATA (2006). Iata website. accessed online. pages 79

INSNA (2008). International networks for social network analysis. accessed online on june, www.insna.org. pages 105

Jafari, J. (2000). *Encyclopedia of Tourism*. Routledge, London, 1 edition. pages 109

James, R. (1997). Hr megatrends. *Human Resource Megatrends*, 36(4):453462. pages 40

Jergensen, F. and Solvoll, G. (1996). Demand models for inclusive tour charter: the norwegian case. *Tourism Management*, 17:17–24. pages 182

Johnson, G. and Scholes, K. (1997). *Exploring corporate strategy*. Prentice Hall, London, 4th edition. pages 39

Jones, C., Hesterly, W. S., and Borgatti, S. P. (1997). A general theory of network governance: Exchange conditions and social mechanisms. *Academy of Management Review*, 22(4):911–945. pages 104, 106

Karlberg, M. (1997). Testing transitivity in graphs. *Soc. Networks*, 19:325–343. pages 139

Kim, W. G., Lee, C., and Hiemstra, S. J. (2004). Effects of an online virtual community on customer loyalty and travel product purchases. *Tourism Management*, 25(3):343–355. pages 35

Knoke, D. and Kuklinski, J. (1983). *Network analysis*. Sage, Los Angeles. pages 83, 84

Krackhardt, D. and Porter, L. W. (1985). When friends leave: A structural analysis of relationship between turnover and stayer's attitudes. *Administrative Science Quarterly*, 30:242–261. pages 104

Krackhardt, D. and Porter, L. W. (1986). The snowball effect: Turnover embedded in communication networks. *Applied Psychology*, 71:50–55. pages 104

Kramer, R., Modsching, M., and Hagen, K. (2005). Development and evaluation of a context-driven, mobile tourist guide. *J. Pervasive Comput. & Comm.*, 1(2):1–12. pages 91, 92

Krapivsky, P. L., Rodgers, G. J., and Redner, S. (2001). Degree distributions of growing networks. *Physical Review Letters*, 86:5401. pages 69

Kuhn, T. S. (1962[1970]). *The Structure of Scientific Revolutions*. University Of Chicago Press. pages 103

Kumar, R., Raghavan, P., Rajalopagan, S., and Tomkins, A. (1999). *Proceedings of the 9th ACM Symposium on Principles of Database Systems*, 1. pages 73

Lang, T. C. (2000). The effect of the internet on travel consumer purchasing behaviour and implications for travel agencies. *Journal of Vacation Marketing*, 6(4):368–385. pages 90

Larsen, J., Urry, J., and Axhausen, K. (2006). Mobilities, networks, geographies. *Growth and Change,*, 39(1):181–184. pages 38

Larson, M. (2002). A political approach to relationship marketing: Case study of the storsjöyran festival. *International Journal of Tourism Research*, 4:119–143. pages 85

Latora, V. and Marchiori, M. (2002). Is the boston subway a small-world network? *Physica A: Statistical Mechanics and its Applications*, 314(1-4):109–113. pages 73

Lawrence, S. and Giles, C. L. (1998). Searching the World Wide Web. *Science*, 280(5360):98–100. pages 73

Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. *Journal of Abnormal and Social Psychology*, 46:38–50. pages 104

Lee, E. J., Goh, K. I., Kahng, B., and Kim, D. (2005). Robustness of the avalanche dynamics in data packet transport on scale-free networks. *Physical Review E*, 71:056108. pages 60

Lemmetyinen, A. and Go, F. M. (2008). The key capabilities required for managing tourism business networks. *Tourism Management*, pages 1–10. pages 82

Lepp, A. and Gibson, H. (2003). Tourist roles, perceived risk and international tourism. *Annals of Tourism Research*, 30:606–624. pages 40

Lew, A. and McKercher, B. (2006). Modeling tourist movement: a local destination analysis. *Annals of Tourism Research*, 33(2):403–423. pages 186

Lewin, K. (1936). *Principles of Topological Psychology*. McGraw-Hill, New York. pages 74, 75

Li, W. and Cai, X. (2003). Statistical analysis of airport network of china. pages 79

Liljeros, F. (2004). Sexual networks in contemporary western societies. *Physics A*, 338:238–245. pages 78

Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H., and Aberg, Y. (2001). The web of human sexual contacts. *Nature*, 411:907. pages 78

Litvin, S. W., Goldsmith, R. E., and Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, 29(3):458–468. pages 38

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20:130–141. pages 106

Lorrain, F. and White, H. (1971). Structural equivalence of infividuals in social networks. *Journal of Mathematical Sociology*, 1:49–80. pages 105

Loudhouse-Research (2007). The american express hospitality monitor. research. pages 35, 37, 151

Lumsdon, L. (2006). Factors affecting the design of tourism bus servics. *Annals of Tourism Research*, 33(3):748766. pages 91

Lusseau, D. and Newman, M. E. J. (2004). Identifying the role that individual animals play in their social network. *Proc. R. Soc. London B (Suppl.)*, 271:S477–S481. pages 76

Lynch, P. and Tinsleyand, R. (2001). Small tourism business networks and destination development. *International Journal of Hospitality Management*, 20(4):367–378. pages 86

Lynch, P. A. (2000). Networking in the homestay sector. *Services Industries Journal*, 20(3):95–116. pages 84, 85, 86

Marsden, P. V. (1982). *Brokerage Behavior in Restricted Exchange Networks*. Sage Publications, Bevery Hills, CA, social structture and network analysis edition. pages 56

Marsden, P. V. and Campbell, K. E. (1984). Measuring tie strength. *Social Forces*, 63(2):482–501. pages 54

Martinez, N. D. (1991). Artifacts or attributes? effects of resolution on the little rock lake food web. *Ecol. Monogr.*, 61:367–392. pages 164

Maslov, S., Sneppen, K., and Zaliznyak, A. (2004). Detectionnext of topological patterns in complex networks: correlation profile of the internet. *Physica A*, 333:529–540. pages 164

Mathews, V. E. (2000). Competition in the international hotel industry. *International Journal of Contemporary Hospitality Management*, 12:114–118(5). pages 92

McDonald, K. (1999). *The Quick Python Book*. Manning Publications Co., Greenwich, CT, USA. pages 112

McGehee, N. and Santos, C. (2005). Social change, discourse and volunteer tourism. *Annals of Tourism Research*, 32(3):760779. pages 82, 87

McKercher, B. (2005). A case for ranking tourism journals. *Tourism Management*, 26:649651. pages 18, 110, 111

McKercher, B. and Fu, C. (2006). Living on the edge. *Annals of Tourism Research*, 33(2):508524. pages 83

McKercher, B., Law, R., and Lam, T. (2005). Rating tourism and hospitality journals. *Tourism Management*, 27:12351252. pages 18, 110, 111

Mehra, A., Kilduff, M., and Brass, D. J. (2001). The social networks of high and low self-monitors: Implications for workplace performance. *Administrative Science Quarterly*, 46:121–146. pages 104

Meyer, J. W. and Rowan, B. (1977). Institutional organizations: formal structure as myth and ceremony. *American Journal of Sociology*, 83:929–984. pages 96

Miguéns, J. and Corfu, A. (2008). e-destination structure: a network analysis approach. Number 17, pages 580–591. SpringerWienNewUork. pages 189

Miguéns, J. I. L. and Mendes, J. F. F. (2008a). Nonsocial patterns of travellers. unpublished. pages 162

Miguéns, J. I. L. and Mendes, J. F. F. (2008b). Travel and tourism: Into a complex network. *Physica A*, 387(12):29632971. pages 93, 94, 151, 152, 163, 164, 173, 186

Miguéns, J. I. L., Mendes, J. F. F., and Costa, C. M. M. (2007). International tourism network. In SPIE, editor, *Proc. SPIE Int. Soc. Opt. Eng.*, volume 6601. pages 149

Molloy, M. and Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structure and Algorithms*, 6:161. pages 64

Molloy, M. and Reed, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability and Computing*, 7:295. pages 64

Monasson, R. (1999). Diffusion, localization and dispersion relations on “small-world” lattices. *Eur. Phys. J. B.*, 12:555. pages 65

Moreno, J. L. (1953,1978). *Who Shall Survive? Foundations of Sociometry, Group Psychotherapy and Sociodrama.* Beacon House, N. Y. Beacon. pages 74

Morrison, A., Lynch, P., and Johns, N. (2004). International tourism networks. *International Journal of Contemporary Hospitality Management*, 16:197–202(6). pages 81, 82, 85

Mossberg, L. L. (1995). Tour leaders and their importance in charter tours. *Tourism Management*, 16:437–445. pages 182

Murphy, L. (2001). Exploring social interactions of backpackers. *Annals of Tourism Research*, 28(1):50–67. pages 83

Nations, U. (1998). Kyoto protocol to the united nations framework convention on global change. report. http://unfccc.int/resource/docs/convkp/kpeng.html, retrived on 20/02/2008. pages 31

Newcomb, S. (1881). Note on the frequency of the use of digits in natural number. *American Journal of Mathematics*, 4:39–40. pages 133

Newman, M. E. J. (2000). Small worlds: The structure of social networks. working paper. pages 48, 104

Newman, M. E. J. (2001). Scientific collaboration networks. i. network construction and fundamental results. *Phys Rev E Stat Nonlin Soft Matter Phys*, 64. pages 60

Newman, M. E. J. (2001 b). The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA*, 98:404–409. pages 54, 76

Newman, M. E. J. (2003a). Assortative mixing in networks. *Phys. Rev. E*, 67. pages 77, 78, 107, 157, 184

Newman, M. E. J. (2003b). The structure and function of complex networks. *SIAM Rev.*, 45:167–256. pages 55, 66, 69

Newman, M. E. J., Forrest, S., and Balthrop, J. (2002). Email networks and the spread of computer viruses. *Phys. Rev. E*, 66:035101. pages 45, 76, 164

Newman, M. E. J. and Park, J. (2003). Why social networks are different from other types of networks. *Phys Rev E Stat Nonlin Soft Matter Phys*, 68(3 Pt 2). pages 163

Newman, M. E. J. and Watts, D. J. (1999). Renormalization group analysis of the small-world network mode. *Phys. Lett. A*, 263:341. pages 59, 65, 66

Nigrini, N. (1996). A taxpayer compliance application of benford's law. *Journal of the American Taxation Association*, 18:72–91. pages 134

Novelli, M., Schmitz, B., and Spencer, T. (2005). Networks, clusters and innovation in tourism: A uk experience. *Tourism Management*, 27:11411152. pages 89, 90

O'Brien, D. (2006). Event business leveraging the sydney 2000 olympic games. *Annals of Tourism Research*, 33(1):240261. pages 88

O'Connor, P. (1999). *Electronic information distribution in tourism and hospitality*. Wallingford: CAB. pages 38

Onnela, J. P., Saramaki, J., Kertesz, J., and Kaski, K. (2005). Intensity and coherence of motifs in weighted complex networks. *Physical Review E*, 71:065103. pages 156, 164, 166, 168

Palmer, A. (1996). Linking external and internal relationship building in networks of public and private sector organisations: a case study. *International Journal of Public Sector Management*, 9(3):51–60. pages 87

Pareto, V. (1896). *Cours d'Economie Politique*. Droz, Genève. pages 114

Park, S. M. and Kim, B. J. (2006). Dynamic behaviors in directed networks. *Physical Review E*, 74:026114. pages 155

Pavlovich, K. (April 2003). The evolution and transformation of a tourism destination network: the waitomo caves, new zealand. *Tourism Management*, 24:203–216(14). pages 82, 86, 93, 94, 95

Pechlaner, H., Zehrer, A., Matzler, K., and Abfalter, D. (2004). A ranking of international tourism and hospitality journals. *Journal of Travel Research*, 42:328–332. pages 18, 110, 111

Petrillo, C. S. and Swarbrooke, J. (2005). *Networking and Partnership in Destination Development and Management*. Napoli. pages 85

Pforr, C. (2006). Tourism policy in the making. an australian network study. *Annals of Tourism Research*, 33:87–108. pages 44, 88

Pietronero, L., Tosatti, E., Tosatti, V., and Vespignani, A. (2001). Explaining the uneven distribution of numbers in nature: The laws of benford and zipf. *Physica*, 304:293–297. pages 134

Poon, A. (1993). Tourism, technology and competitive strategies. Technical report, Wallingford: CAB International. pages 35, 97

Portes, A. (1998). Social capital: its origins and applications in modern sociology. *Annual Review of Sociology*, 24:1–24. pages 57, 105

Powell, W. W., Koput, K. W., and Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology. *Administrative Science Quartely*, 41:116–145. pages 95

Price, D. J. D. (1963). *Little Science, big Science.* New York: Columbia University Press. pages 103

Radcliffe-Brown, A. R. (1930). The social organization of australian tribes. *Oceania Monograph*, 1:34–63. pages 75

Ram, T. M. (1994). A ricardian analysis of the fully inclusive tour industry. *Service Industries Journal*, 14(1):85–93. pages 86

Renzulli, L. A., Aldrich, H. E., and Moody, J. (2000). Family matters: Gender, networks, and entrepreneurial outcomes. *Social Forces*, 79:523–546. pages 104

Ritchie, J. (1999). Crafting a value-driven vision. *Tourism Management*, 20:273–282. pages 93

Rosenfeld, S. A. (2001). Backing into clusters: Retrofitting public policies. report: John f. kennedy school symposium, Harvard University. pages 83

Rowley, T. J. (1997). Moving beyond dyadic ties: a network theory of stakeholder influence. *Academy of Management Review*, 22(4):887–910. pages 95, 96, 104

Ryan, C. (2005). The ranking and rating of academics and journals in tourism research. *Tourism Management*, 26:653656. pages 18, 110, 111

Sandwith, T., Shine, C., Hamilton, L., and Sheppard, D. (2001). Transboundary protected areas for peace and co-operation. Technical report, IUCN, Gland, Switzerland and Cambridge, UK. pages 89

Saramaki, J., Kivela, M., Onnela, J. P., Kaski, K., and Kertesz, J. (2007). Generalizations of the clustering coefficient to weighted complex networks. *Phys. Rev. E*, 75:027105. pages 156

Saxena, G. (2005). Relationships, networks and the learning regions:case evidence from the peak district national park. *Tourism Management*, 26:277289. pages 82

Schelling, T. C. (1978). *Micromotives and Macrobehavior*. New York: W. W. Norton. pages 82

Sciarini, P. (1996). Elaboration of the swiss agricultural policy for the gatt negotiations: A network analysis. *Schweizer Zeitschrift für Soziologie*, 22:85–115. pages 106

Scott, J. P. (2000). *Social Network Analysis: A Handbook*. SAGE Publications. pages 38, 44, 53, 55, 56, 74, 76, 95, 107, 109

Scott, N., Baggio, R., and Cooper, C. (2008a). *Network Analysis and Tourism: From Theory to Practice (Aspects of Tourism): From Theory to Practice*. Channel View Publications. pages 85, 94

Scott, N., Cooper, C., and Baggio, R. (2007). *Proceedings of the 2007 advances in tourism marketing conference [CD-ROM]*, chapter Use of network analysis in tourism. Valencia, Spain: University of Valencia. pages 93

Scott, N., Cooper, C., and Baggio, R. (2008b). Destination networks four australian cases. *Annals of Tourism Research*, 35(1):169–188. pages 81, 85, 87, 93, 95, 96, 97, 185, 187

Seidman, S. B. (1983). Network structure and minimum degree. *Social Networks*, 5:269–287. pages 61, 171

Serrano, M. A., Boguna, M., and Vespignani, A. (2007). Patterns of dominant flows in the world trade web. *J. Econ. Interac. Coor.*, 2:111. pages 151

Shaw, E. and Conway, S. (2000). *Networking and the small firm*. Finantial Times/Prentice Hall, Harlow. pages 82, 84

Sheet, T. F. (2008). http://www.tripadvisor.com/presscenter-c4-fact_sheet.html. online. pages 38

Sheldon, P. (1998). Tourism information technology. *Wallingford: CAB International*. pages 35

Shih, H. Y. (2006). Network characteristics of drive tourism destinations: an application of network analysis in tourism. *Tourism Management*, 27(5):1029–1039. pages 57, 92, 93, 95, 96, 97, 185, 187, 190, 194

Sigaux, G. (1876). *The History of Tourism*. Edito Service Ltd., Geneva, Switzerland. pages 29

Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika*, 42:425–440. pages 69

Smeral, E. (1998). The impact of globalization on small and medium enterprises: new challenges for tourism policies in european countries. *Tourism Management*, 19:371–380. pages 90

Smith, K. (1993). The influence of weather and climate on recreation and tourism. *Weather*, 48(12):398–404. pages 28, 31, 104

Smith-Ring, P. (1999). *The formation of inter-organizational networks*, chapter Processes facilitating reliance on trust in interorganizational networks, pages 113–145. Oxford University Press, Oxford. pages 84

Solomonoff, R. and Rapoport, A. (1951). Connectivity of random nets. *Bulletin of Mathematical Biophysics*, 13:107. pages 62

Sparrowe, R. T. and Liden, R. C. (1997). Process and structure in leader-member exchange. *Academy of Mangement Review*, 22:193–234. pages 104

Stiller, J., Nettle, D., and Dunbar, R. (2003). The small world of shakespeares plays. *Human Nature*, 14:397. pages 76

Strogatz, H. (2001). Exploring complex networks. *Nature*, 410:268–276. pages 45, 58, 73, 104, 175

Szarka, J. (1990). Networking and small firms. *International Small Business Journal*, 8(2):10–21. pages 83

Tadic, B. (2001). Access time of an adaptive random walk on the world-wide-web. *Physica A*, 293:273. pages 69

Third-World-Network (2006). Global development finance 2001. Technical report, World Bank report, accessed on 01.10.2008 http://www.twnside.org.sg/title/forecasts.htm. pages 39

Thompson, G., Frances, J., Levacic, R., and Mitchell, J. (1991). *Markets, Hierarchies & Networks: The Coordination of Social Life*. Sage, London. pages 83, 84

Travers, J. and Milgram, S. (1969). An experimental study of the small world problem. *Sociometry*, 32(4):425–443. pages 23, 49, 76, 104

Tufts, S. and Milne, S. (1999). Museums: A supply-side perspective. *Annals of Tourism Research*, 26:613–631. pages 87

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, 42(1):35–67. pages 96

Valente, T. W. and Foreman, R. K. (1998). Integration and radiality: Measuring the extent of an individual's connectedness and reachability in a network. *Social Networks*, 20:85–109. pages 76

Vázquez, A., Pastor-Satorras, R., and Vespignani, A. (2002). Large-scale topological and dynamical properties of the internet. *Phys. Rev. E*, 65:66130. pages 74

Wachi, S., Yoneda, K., and Wu, R. (2005). Interactone-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. *Bioinformatics*, 21(23):4205–4208. pages 61

Wang, Y. and Fesenmaier, D. (2002). Defining the virtual tourist community: implications for tourism marketing. *Tourism Management*, 23(4):407–417. pages 35

Wang, Y. and Fesenmaier, D. (2007). Collaborative destination marketing: A case study of elkhart county, indiana. *Tourism Management*, 28(3):863–875. pages 85, 89

Wasserman, S., Faust, K., and Iacobucci, D. (1994). *Social Network Analysis : Methods and Applications (Structural Analysis in the Social Sciences)*. Cambridge University Press. pages 38, 44, 53, 55, 58, 74, 75, 76, 95, 103, 104, 106, 108, 109, 138, 183

Watts, D. J. (2003a). *Six Degrees: The Science of a Connected Age*. W. W. Norton & Company. pages 38, 48

Watts, D. J. (2003b). *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton University Press. pages 44, 48, 49, 50, 55, 63, 109

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442. pages 50, 63, 159, 173

Wellman, B. (1988). *Social structures: A network approach*, chapter Structural analysis: From method and metaphor to theory and substance, pages 19–61. Cambridge: Cambridge University Press. pages 104, 106

Werde, A., Tampubolon, G., and Savage, M. (2005). Recreation, informal social networks and social capital. *Journal of Leisure Research*, 37(4):402–425. pages 93

West, D. B. (1995). *Introduction to Graph Theory*. Prentice-Hall, Englewood Cliffs NJ. pages 63

White, D. R. and Reitz, K. P. (1985). Graph and semigroup homomorphisms on networks of relations. *Social Networks*, 5:193–234. pages 105

White, H. C., Boorman, S. A., and Breiger, R. L. (1976). Social-structure from multiple networks i: Blockmodels of roles and positions. *American Journal of Sociology*, 81:730–780. pages 75

Willmott, M. and Graham, S. (2001). *Tourism and hospitality in the 21st century*, chapter The world of today and tomorrow: the European picture, page 2938. Butterworth-Heinemann, Oxford. pages 40

Wolfram, S. (2003). *The Mathematica Book, Fifth Edition*. Wolfram Media. pages 112

Wright, P. (1996). *Beginning Visual Basic 5*. Wrox Press Ltd., Birmingham, UK, UK. pages 112

WTO (2000). *Global Tourism Forecast to the Year 2000 and Beyond: The World*. World Tourism Organization Pbcn, Madrid. pages 90, 117

WTO, editor (2004). *Long-term Prospects: Tourism 2020 Vision*. World Tourism Organization. pages 24, 27, 28, 32

WTO (2006). *New Yearbook of Tourism Statistics Pubns*. World Tourism Organization Pbcn. pages 118, 123, 126, 127, 136

WTO, editor (2007). *UNWTO Tourism Highlights, Edition 2007*. World Tourism Organization. pages 27, 32, 34, 40

Wuchty, S. and Almaas, E. (2005). Peeling the yeast protein network. *Proteomics*, 5:444–449. pages 45, 61

Yan, G., Zhou, T., Wang, J., Fu, Z., and Wang, B. (2005). Empirical study on chinese city airport networks. *Chinese Phys. Lett.*, 22:510. pages 70

Yuksel, A. and Yuksel, F. (2004). Managing relations in a learning model for bringing destinations in need of assistance into contact with good practice. *Tourism Management*, 26:667–679. pages 82, 85

Zach, F., Gretzel, U., and Fesenmaier, D. (2008). *Information and Communication Technologies in Tourism 2008*, proceedings of the international conference in innsbruck Tourist activated networks: implications for dynamic packaging systems in tourism, pages 198–208. Springer-WienNewYork, Wien. pages 20, 61, 96, 97, 185, 190

Zhang, B. and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology*, 4(17). pages 156

Zhou, S. and Mondragón, R. J. (2004). Accurately modeling the internet topology. *Phys. Rev. E*, 70(6):066108. pages 80, 160, 173

Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort*. Addison-Wesley (Reading MA). pages 114